

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2004, 14:22:29 ; Search time 26.3838 Seconds
(without alignments)
3095.051 Million cell updates/sec

Title: US-10-042-431-49

Perfect score: 2076

Sequence: 1 VHMTKAVDPFAFNISEII.....IHLMOBETNLSQRCRAVL 390

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubaa/FCI_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2076	100.0	390	11	US-09-759-130B-419
2	2076	100.0	390	14	US-10-042-431-49
3	2076	100.0	409	10	US-09-835-996A-21
4	2076	100.0	423	11	US-09-759-130B-417
5	2076	100.0	423	14	US-10-042-431-47
6	1991	95.9	374	10	US-09-811-825-4
7	1991	95.9	398	11	US-09-811-825-2
8	1289	62.1	399	11	US-09-759-130B-411
9	1289	62.1	399	12	US-10-051-874-92
10	1289	62.1	399	12	US-10-051-874-93
11	1289	62.1	399	12	US-10-051-874-94
12	1289	62.1	399	14	US-10-042-431-41
13	1282	61.8	392	14	US-10-003-302-4
14	1165	56.1	221	11	US-09-759-130B-420
15	1165	56.1	221	14	US-10-042-431-50

16	1161	55.9	395	12	US-10-015-115-98
17	1141.5	55.0	398	12	US-10-015-115-96
18	1141	55.0	357	12	US-10-108-260A-3705
19	1129	54.4	399	12	US-10-221-097-35
20	1127	54.3	378	12	US-10-420-564-2
21	1126	54.2	397	12	US-10-312-088-29
22	1109	53.4	371	12	US-10-015-115-97
23	1109	53.4	398	11	US-09-759-130B-445
24	1109	53.4	398	12	US-10-015-115-94
25	1109	53.4	398	12	US-10-015-115-95
26	1109	53.4	398	14	US-10-042-431-75
27	1082	52.1	398	14	US-10-056-744B-2
28	1024	49.3	401	12	US-10-015-115-32
29	1023	49.3	395	14	US-10-003-302-2
30	998	48.1	373	12	US-10-051-874-28
31	995	47.9	373	12	US-10-051-874-30
32	989	46.7	365	12	US-10-312-088-28
33	989	46.7	395	12	US-10-051-874-90
34	929	44.7	349	12	US-10-051-874-32
35	890.5	42.9	390	12	US-10-015-115-30
36	835	40.2	351	12	US-10-051-874-91
37	783	37.7	144	11	US-09-759-130B-422
38	783	37.7	144	14	US-10-042-431-52
39	400	19.3	127	14	US-10-056-744B-5
40	319	15.4	567	12	US-10-032-585-7720
41	296	14.3	476	15	US-10-128-714-3075
42	296	14.3	641	15	US-10-128-714-3075
43	294	14.2	431	12	US-10-304-038-6
44	209	10.1	86	12	US-09-864-408A-8008
45	128	6.2	25	11	US-09-759-130B-421

ALIGNMENTS

RESULT 1

US-09-759-130B-419
; Sequence 419, Application US/09759130B
; Publication No. US2003002279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A
; APPLICANT: Fraser, Christopher C
; APPLICANT: Sharp, John D
; APPLICANT: Barnes, Thomas S
; APPLICANT: Kirt, Susan J
; APPLICANT: Mackay, Charles R
; APPLICANT: Myers, Paul S
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; TITLE OF INVENTION: USES.
; FILE REFERENCE: MPI00-5350NMIM
; CURRENT APPLICATION NUMBER: US/09/759,130B
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996

; PRIOR FILING DATE: 1999-09-10
 ; PRIOR APPLICATION NUMBER: US 09/602,871
 ; PRIOR FILING DATE: 2000-06-23
 ; PRIOR APPLICATION NUMBER: US 09/420,707
 ; PRIOR FILING DATE: 1999-10-19
 ; NUMBER OF SEQ ID NOS: 460
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 419
 ; LENGTH: 390
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-759-130B-419

Query Match 100.0%; Score 2076; DB 11; Length 390;
 Best Local Similarity 100.0%; Pred. No. 7.3e-213;
 Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VHMPTKAVDPEAFNWISEIIHQGYPCPEYEVATEDGYILSVNRPGLVQPKTGSRPV 60
 DB 1 VHMPTKAVDPEAFNWISEIIHQGYPCPEYEVATEDGYILSVNRPGLVQPKTGSRPV 60
 QY 61 VLLQHLVGGASNWIENLNNPNSLGFILADAGPDVWNGSNGNAWSKHKHTLSIDQDEFWA 120
 DB 61 VLLQHLVGGASNWIENLNNPNSLGFILADAGPDVWNGSNGNAWSKHKHTLSIDQDEFWA 120
 QY 121 FSYDEMARFDPVAVINFILOKTCQEKIYVGVSGTGMGFIAPSTMPPELAQKIKMYFALA 180
 DB 121 FSYDEMARFDPVAVINFILOKTCQEKIYVGVSGTGMGFIAPSTMPPELAQKIKMYFALA 180
 QY 181 PIATVHKASPGTKFLLPDMIMKGLFGKKEFLYQTRFLRQLVYLCGOVILDDQICSNIM 240
 DB 181 PIATVHKASPGTKFLLPDMIMKGLFGKKEFLYQTRFLRQLVYLCGOVILDDQICSNIM 240
 QY 241 LLLGGFNTNNMNSRASVYAAHTLAGTSVQNIHLWSQAVNSGELRAFDWGSSETKNLEKCN 300
 DB 241 LLLGGFNTNNMNSRASVYAAHTLAGTSVQNIHLWSQAVNSGELRAFDWGSSETKNLEKCN 300
 QY 301 OPTPVRYRVDMTVPTAMWTGGODWLSNPEDVKMLISEVTNLIYHKNIPWAHVDFIWL 360
 DB 301 OPTPVRYRVDMTVPTAMWTGGODWLSNPEDVKMLISEVTNLIYHKNIPWAHVDFIWL 360
 QY 361 DAPHRMYNEIHLMQQBEETNLSQGRCEAVL 390
 DB 361 DAPHRMYNEIHLMQQBEETNLSQGRCEAVL 390

RESULT 2
 US-10-042-431-49
 ; Sequence 49, Application US/10042431
 ; Publication No. US20020182675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MCCARTHY, Sean A
 ; APPLICANT: BARNES, Thomas M
 ; APPLICANT: FRASER, Christopher C
 ; APPLICANT: SHARP, John D
 ; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
 ; PREVENTIVE, THERAPEUTIC, AND OTHER USES
 ; FILE REFERENCE: 10147-602
 ; CURRENT APPLICATION NUMBER: US/10/042,431
 ; CURRENT FILING DATE: 2001-10-25
 ; PRIOR APPLICATION NUMBER: US 09/333,159
 ; PRIOR FILING DATE: 1999-06-14
 ; PRIOR APPLICATION NUMBER: US 09/578,063
 ; PRIOR FILING DATE: 2000-05-24
 ; NUMBER OF SEQ ID NOS: 79
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 49
 ; LENGTH: 390
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-042-431-49
 Query Match 100.0%; Score 2076; DB 14; Length 390;

Best Local Similarity 100.0%; Pred. No. 7.3e-213;
 Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VHMPTKAVDPEAFNWISEIIHQGYPCPEYEVATEDGYILSVNRPGLVQPKTGSRPV 60
 DB 1 VHMPTKAVDPEAFNWISEIIHQGYPCPEYEVATEDGYILSVNRPGLVQPKTGSRPV 60
 QY 61 VLLQHLVGGASNWIENLNNPNSLGFILADAGPDVWNGSNGNAWSKHKHTLSIDQDEFWA 120
 DB 61 VLLQHLVGGASNWIENLNNPNSLGFILADAGPDVWNGSNGNAWSKHKHTLSIDQDEFWA 120
 QY 121 FSYDEMARFDPVAVINFILOKTCQEKIYVGVSGTGMGFIAPSTMPPELAQKIKMYFALA 180
 DB 121 FSYDEMARFDPVAVINFILOKTCQEKIYVGVSGTGMGFIAPSTMPPELAQKIKMYFALA 180
 QY 181 PIATVHKASPGTKFLLPDMIMKGLFGKKEFLYQTRFLRQLVYLCGOVILDDQICSNIM 240
 DB 181 PIATVHKASPGTKFLLPDMIMKGLFGKKEFLYQTRFLRQLVYLCGOVILDDQICSNIM 240
 QY 241 LLLGGFNTNNMNSRASVYAAHTLAGTSVQNIHLWSQAVNSGELRAFDWGSSETKNLEKCN 300
 DB 241 LLLGGFNTNNMNSRASVYAAHTLAGTSVQNIHLWSQAVNSGELRAFDWGSSETKNLEKCN 300
 QY 301 OPTPVRYRVDMTVPTAMWTGGODWLSNPEDVKMLISEVTNLIYHKNIPWAHVDFIWL 360
 DB 301 OPTPVRYRVDMTVPTAMWTGGODWLSNPEDVKMLISEVTNLIYHKNIPWAHVDFIWL 360
 QY 361 DAPHRMYNEIHLMQQBEETNLSQGRCEAVL 390
 DB 361 DAPHRMYNEIHLMQQBEETNLSQGRCEAVL 390

RESULT 3
 US-09-835-996A-21
 ; Sequence 21, Application US/09835996A
 ; Patent No. US20020142953A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ballinger, Dennis
 ; APPLICANT: Loeb, Debra
 ; APPLICANT: Montgomery, Julie
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Zhou, Ping
 ; APPLICANT: Goodrich, Ryle
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Zhao, Qing
 ; APPLICANT: Wehrman, Tom
 ; APPLICANT: Drmanac, Radoje
 ; APPLICANT: Ren, Feiyan
 ; APPLICANT: Qian, Xiaohong
 ; APPLICANT: Wang, Dunrui
 ; TITLE OF INVENTION: MATERIALS AND METHODS RELATING TO LIPID METABOLISM
 ; FILE REFERENCE: 28110/35915A
 ; CURRENT APPLICATION NUMBER: US/09/835,996A
 ; CURRENT FILING DATE: 2001-04-16
 ; PRIOR APPLICATION NUMBER: US 60/197,137
 ; PRIOR FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: US 09/714,936
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: US 09/667,298
 ; PRIOR FILING DATE: 2000-09-22
 ; PRIOR APPLICATION NUMBER: US 09/631,451
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: US 09/598,042
 ; PRIOR FILING DATE: 2000-06-20
 ; NUMBER OF SEQ ID NOS: 45
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 21
 ; LENGTH: 409
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-835-996A-21

Query Match 100.0%; Score 2076; DB 10; Length 409;
Best Local Similarity 100.0%; Pred. No. 7.8e-213;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VHMPTKAVDPPEAFNMISIIHQHQPCEYEYVATEDGYILSVNRIIPRGLVQPKKTGSRPV 60
Db 20 VHMPTKAVDPPEAFNMISIIHQHQPCEYEYVATEDGYILSVNRIIPRGLVQPKKTGSRPV 79
QY 61 VLLQHLVGGASNWI SNL PNNNSLGFILADAGFDVVMGNSRGNNAWSRKHKTLSIDQDEFWA 120
Db 80 VLLQHLVGGASNWI SNL PNNNSLGFILADAGFDVVMGNSRGNNAWSRKHKTLSIDQDEFWA 139
QY 121 PSYDEMARFDPVAVINFILOKTGQEKIYYVGSQGTMGFIATFSTMPPELAQIKMYFALA 180
Db 140 PSYDEMARFDPVAVINFILOKTGQEKIYYVGSQGTMGFIATFSTMPPELAQIKMYFALA 199
QY 181 PIATVKHAKSPCTKELLPLDMWIKGLFGKKEFLYQTRFLQVLYLCCQVILDOICSNIM 240
Db 200 PIATVKHAKSPCTKELLPLDMWIKGLFGKKEFLYQTRFLQVLYLCCQVILDOICSNIM 259
QY 241 LLLGGFNTNNMNSRASVYAAHTLAGTSVQNILHWSQAVNSGELRAFDFWGSSETKNLEKCN 300
Db 260 LLLGGFNTNNMNSRASVYAAHTLAGTSVQNILHWSQAVNSGELRAFDFWGSSETKNLEKCN 319
QY 301 QPTPVRYRVRDMTVP TAMWTGGQDWLSNPEDVKMLLSVNTNLIYHKNIPEWAHVDFIWL 360
Db 320 QPTPVRYRVRDMTVP TAMWTGGQDWLSNPEDVKMLLSVNTNLIYHKNIPEWAHVDFIWL 379
QY 361 DAPHRMYNEIHLMOQEBETNLSQGRCEAVL 390
Db 380 DAPHRMYNEIHLMOQEBETNLSQGRCEAVL 409

RESULT 4

US-09-759-130B-417
; Sequence 417, Application US/09759130B
; Publication No. US20030022279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A
; APPLICANT: Fraser, Christopher C
; APPLICANT: Sharp, John D
; APPLICANT: Barnes, Thomas S
; APPLICANT: KIRST, Susan J
; APPLICANT: Mackay, Charles R
; APPLICANT: Myers, Paul S
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; FILE REFERENCE: MF100-5350MIM
; CURRENT APPLICATION NUMBER: US/097591,130B
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/602,871

; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/420,707
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 417
; LENGTH: 423
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-759-130B-417

Query Match 100.0%; Score 2076; DB 11; Length 423;
Best Local Similarity 100.0%; Pred. No. 8.2e-213;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VHMPTKAVDPPEAFNMISIIHQHQPCEYEYVATEDGYILSVNRIIPRGLVQPKKTGSRPV 60
Db 34 VHMPTKAVDPPEAFNMISIIHQHQPCEYEYVATEDGYILSVNRIIPRGLVQPKKTGSRPV 93
QY 61 VLLQHLVGGASNWI SNL PNNNSLGFILADAGFDVVMGNSRGNNAWSRKHKTLSIDQDEFWA 120
Db 94 VLLQHLVGGASNWI SNL PNNNSLGFILADAGFDVVMGNSRGNNAWSRKHKTLSIDQDEFWA 153
QY 121 PSYDEMARFDPVAVINFILOKTGQEKIYYVGSQGTMGFIATFSTMPPELAQIKMYFALA 180
Db 154 PSYDEMARFDPVAVINFILOKTGQEKIYYVGSQGTMGFIATFSTMPPELAQIKMYFALA 213
QY 181 PIATVKHAKSPCTKELLPLDMWIKGLFGKKEFLYQTRFLQVLYLCCQVILDOICSNIM 240
Db 214 PIATVKHAKSPCTKELLPLDMWIKGLFGKKEFLYQTRFLQVLYLCCQVILDOICSNIM 273
QY 241 LLLGGFNTNNMNSRASVYAAHTLAGTSVQNILHWSQAVNSGELRAFDFWGSSETKNLEKCN 300
Db 274 LLLGGFNTNNMNSRASVYAAHTLAGTSVQNILHWSQAVNSGELRAFDFWGSSETKNLEKCN 333
QY 301 QPTPVRYRVRDMTVP TAMWTGGQDWLSNPEDVKMLLSVNTNLIYHKNIPEWAHVDFIWL 360
Db 334 QPTPVRYRVRDMTVP TAMWTGGQDWLSNPEDVKMLLSVNTNLIYHKNIPEWAHVDFIWL 393
QY 361 DAPHRMYNEIHLMOQEBETNLSQGRCEAVL 390
Db 394 DAPHRMYNEIHLMOQEBETNLSQGRCEAVL 423

RESULT 5

US-10-042-431-47
; Sequence 47, Application US/10042431
; Publication No. US20020182675A1
; GENERAL INFORMATION:
; APPLICANT: MCCARTHY, Sean A
; APPLICANT: BARNES, Thomas M
; APPLICANT: FRASER, Christopher C
; APPLICANT: SHARP, John D
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
; TITLE OF INVENTION: PREVENTIVE, THERAPEUTIC, AND OTHER USES
; FILE REFERENCE: 10147-6U2
; CURRENT APPLICATION NUMBER: US/10/042,431
; CURRENT FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 47
; LENGTH: 423
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-042-431-47

Query Match 100.0%; Score 2076; DB 14; Length 423;
Best Local Similarity 100.0%; Pred. No. 8.2e-213;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VMPTKAVDPAPMNISEIIHQGVPCPEYEVATEDGYILSVNRIIPRGIVOPKKTGSRPV 60
 DB 34 VMPTKAVDPAPMNISEIIHQGVPCPEYEVATEDGYILSVNRIIPRGIVOPKKTGSRPV 93
 QY 61 VLLQGLVGGASNWSINLPNNLSGLFADAGDVMGNSRGNAWSRKHTLSIDQDEFWA 120
 DB 94 VLLQGLVGGASNWSINLPNNLSGLFADAGDVMGNSRGNAWSRKHTLSIDQDEFWA 153
 QY 121 FSYDEMAREFDLPAVINFILOKQGOEKIYVGVSGTWTGFIAPSTMPPELAQIKMYFALA 180
 DB 154 FSYDEMAREFDLPAVINFILOKQGOEKIYVGVSGTWTGFIAPSTMPPELAQIKMYFALA 213
 QY 181 PIATVHAQSPGKTFLLPDMIMKGLFGKKEFLYQTRFLRQVLYLCCGVILDDQICSNIM 240
 DB 214 PIATVHAQSPGKTFLLPDMIMKGLFGKKEFLYQTRFLRQVLYLCCGVILDDQICSNIM 273
 QY 241 LLLGGFNTNNMNSRASVVAHTLAGTSVONILHWSQAVNSGELRAFDWSETKNLEKCN 300
 DB 274 LLLGGFNTNNMNSRASVVAHTLAGTSVONILHWSQAVNSGELRAFDWSETKNLEKCN 333
 QY 301 QPTPVYRVRDVTPTAMWTGQDMLSNPDEVKMLLSEVTNLIYHKNIPEWAHVDPIWGL 360
 DB 334 QPTPVYRVRDVTPTAMWTGQDMLSNPDEVKMLLSEVTNLIYHKNIPEWAHVDPIWGL 393
 QY 361 DAPHRMYNEIHLMOQEEETNLSQGRCEAVL 390
 DB 394 DAPHRMYNEIHLMOQEEETNLSQGRCEAVL 423

RESULT 6
 US-09-811-825-4
 ; Sequence 4, Application US/09811825
 ; Patent No. US20020144297A1
 ; GENERAL INFORMATION:
 ; APPLICANT: YAN, Chunhua et al.
 ; TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES
 ; TITLE OF INVENTION: THEREOF
 ; FILE REFERENCE: CL001170
 ; CURRENT APPLICATION NUMBER: US/09/811,825
 ; CURRENT FILING DATE: 2001-03-20
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 374
 ; TYPE: PRT
 ; ORGANISM: Human
 US-09-811-825-4

Query Match 95.9%; Score 1991; DB 10; Length 374;
 Best Local Similarity 100.0%; Pred. No. 8.3e-204;
 Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 SEIIHQGVPCPEYEVATEDGYILSVNRIIPRGIVOPKKTGSRPVVLLQGLVGGASNWS 76
 DB 1 SEIIHQGVPCPEYEVATEDGYILSVNRIIPRGIVOPKKTGSRPVVLLQGLVGGASNWS 60
 QY 77 NLPNNLSGLFADAGDVMGNSRGNAWSRKHTLSIDQDEFWAFSYDEMARFDLPVAIN 136
 DB 61 NLPNNLSGLFADAGDVMGNSRGNAWSRKHTLSIDQDEFWAFSYDEMARFDLPVAIN 120
 QY 137 FILOKQGOEKIYVGVSGTWTGFIAPSTMPPELAQIKMYFALAPIATVYKAKSPGKFL 196
 DB 121 FILOKQGOEKIYVGVSGTWTGFIAPSTMPPELAQIKMYFALAPIATVYKAKSPGKFL 180
 QY 197 LLPDMMIKGLFGKKEFLYQTRFLRQVLYLCCGVILDDQICSNIMLLGGFNTNNMNSRA 256
 DB 181 LLPDMMIKGLFGKKEFLYQTRFLRQVLYLCCGVILDDQICSNIMLLGGFNTNNMNSRA 240
 QY 257 SVYAAHTLAGTSVONILHWSQAVNSGELRAFDWSETKNLEKCNQPTPVYRVRDVTPT 316
 DB 241 SVYAAHTLAGTSVONILHWSQAVNSGELRAFDWSETKNLEKCNQPTPVYRVRDVTPT 300

QY 317 AMWTGGQDMLSNPDEVKMLLSEVTNLIYHKNIPEWAHVDPIWGLDAPHRMYNEIHLMOQ 376
 DB 301 AMWTGGQDMLSNPDEVKMLLSEVTNLIYHKNIPEWAHVDPIWGLDAPHRMYNEIHLMOQ 360
 QY 377 EETNLSQGRCEAVL 390
 DB 361 EETNLSQGRCEAVL 374

RESULT 7
 US-09-811-825-2
 ; Sequence 2, Application US/09811825
 ; Patent No. US20020144297A1
 ; GENERAL INFORMATION:
 ; APPLICANT: YAN, Chunhua et al.
 ; TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES
 ; TITLE OF INVENTION: THEREOF
 ; FILE REFERENCE: CL001170
 ; CURRENT APPLICATION NUMBER: US/09/811,825
 ; CURRENT FILING DATE: 2001-03-20
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 398
 ; TYPE: PRT
 ; ORGANISM: Human
 US-09-811-825-2

Query Match 95.9%; Score 1991; DB 10; Length 398;
 Best Local Similarity 100.0%; Pred. No. 9.1e-204;
 Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 SEIIHQGVPCPEYEVATEDGYILSVNRIIPRGIVOPKKTGSRPVVLLQGLVGGASNWS 76
 DB 25 SEIIHQGVPCPEYEVATEDGYILSVNRIIPRGIVOPKKTGSRPVVLLQGLVGGASNWS 84
 QY 77 NLPNNLSGLFADAGDVMGNSRGNAWSRKHTLSIDQDEFWAFSYDEMARFDLPVAIN 136
 DB 85 NLPNNLSGLFADAGDVMGNSRGNAWSRKHTLSIDQDEFWAFSYDEMARFDLPVAIN 144
 QY 137 FILOKQGOEKIYVGVSGTWTGFIAPSTMPPELAQIKMYFALAPIATVYKAKSPGKFL 196
 DB 145 FILOKQGOEKIYVGVSGTWTGFIAPSTMPPELAQIKMYFALAPIATVYKAKSPGKFL 204
 QY 197 LLPDMMIKGLFGKKEFLYQTRFLRQVLYLCCGVILDDQICSNIMLLGGFNTNNMNSRA 256
 DB 205 LLPDMMIKGLFGKKEFLYQTRFLRQVLYLCCGVILDDQICSNIMLLGGFNTNNMNSRA 264
 QY 257 SVYAAHTLAGTSVONILHWSQAVNSGELRAFDWSETKNLEKCNQPTPVYRVRDVTPT 316
 DB 265 SVYAAHTLAGTSVONILHWSQAVNSGELRAFDWSETKNLEKCNQPTPVYRVRDVTPT 324
 QY 317 AMWTGGQDMLSNPDEVKMLLSEVTNLIYHKNIPEWAHVDPIWGLDAPHRMYNEIHLMOQ 376
 DB 325 AMWTGGQDMLSNPDEVKMLLSEVTNLIYHKNIPEWAHVDPIWGLDAPHRMYNEIHLMOQ 384
 QY 377 EETNLSQGRCEAVL 390
 DB 385 EETNLSQGRCEAVL 398

RESULT 8
 US-09-759-130B-411
 ; Sequence 411, Application US/09759130B
 ; Publication No. US20030022279A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Millennium Pharmaceuticals, Inc.
 ; APPLICANT: McCarthy, Sean A
 ; APPLICANT: Fraser, Christopher C
 ; APPLICANT: Sharp, John D
 ; APPLICANT: Barnes, Thomas S

APPLICANT: Kirt, Susan J
 APPLICANT: Mackay, Charles R
 APPLICANT: Myers, Paul S
 APPLICANT: Leiby, Kevin R
 APPLICANT: Wrighton, Nicolas
 APPLICANT: Gooden, Andrew
 APPLICANT: Holtzman, Douglas A
 TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
 TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
 TITLE OF INVENTION: USES
 FILE REFERENCE: ME100-5350MIM
 CURRENT APPLICATION NUMBER: US/09/759,130B
 CURRENT FILING DATE: 2002-09-16
 PRIOR APPLICATION NUMBER: US 09/479,249
 PRIOR FILING DATE: 2000-01-07
 PRIOR APPLICATION NUMBER: US 09/559,497
 PRIOR FILING DATE: 2000-04-27
 PRIOR APPLICATION NUMBER: US 09/578,063
 PRIOR FILING DATE: 2000-05-24
 PRIOR APPLICATION NUMBER: US 09/333,159
 PRIOR FILING DATE: 1999-06-14
 PRIOR APPLICATION NUMBER: US 09/596,194
 PRIOR FILING DATE: 2000-07-14
 PRIOR APPLICATION NUMBER: US 09/342,364
 PRIOR FILING DATE: 1999-06-29
 PRIOR APPLICATION NUMBER: US 09/608,452
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: US 09/393,996
 PRIOR FILING DATE: 1999-09-10
 PRIOR APPLICATION NUMBER: US 09/602,871
 PRIOR FILING DATE: 2000-06-23
 PRIOR APPLICATION NUMBER: US 09/420,707
 PRIOR FILING DATE: 1999-10-19
 NUMBER OF SEQ ID NOS: 460
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 411
 LENGTH: 399
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-759-130B-411

Query Match 62.1%; Score 1289; DB 11; Length 399;

Best Local Similarity 63.5%; Pred. No. 9,7e-129;

Matches 235; Conservative 55; Mismatches 80; Indels 0; Gaps 0;

Qy	7	AVDPEAFMNTSEIIHQHGYPCCEVEVATEDGYILSVNRIPLGLVQPKTGRSPVVLQHG	66
Db	28	AVDPEAFMNTSEIIHQHGYPCCEVEVATEDGYILSVNRIPLGLVQPKTGRSPVVLQHG	87
Qy	67	LVGGASWISNLPNNSLGFILADGPDVWNGSRGNASRKHKTLSIDQDEFWAFSYDEM	126
Db	88	LVGGASWISNLPNNSLGFILADGPDVWNGSRGNASRKHKTLSIDQDEFWAFSYDEM	147
Qy	127	ARFDLPAVINELLOKTEQEKIYVYSGQTMGFIASTWPELAOKIKYFALAPIATVK	186
Db	148	ARFDLPAVINELLOKTEQEKIYVYSGQTMGFIASTWPELAOKIKYFALAPIATVK	207
Qy	187	HAKSPGTFKLLPDMWIKLFGKKEFLYQTRFLRQLVYLCGVILDOICSNMILLGGF	246
Db	208	HAKSPGTFKLLPDMWIKLFGKKEFLYQTRFLRQLVYLCGVILDOICSNMILLGGF	267
Qy	247	NTNNNWSRASYAAHTLAGTSVONILHWSQVNSGELPAFNGSETNLEKCNQPTTVR	306
Db	268	NTNNNWSRASYAAHTLAGTSVONILHWSQVNSGELPAFNGSETNLEKCNQPTTVR	327
Qy	307	YRVDMTVPTAMTGGQDMLSNPBDVKMLLSEVNTLIYHKNIPWAHVDIFWGLDAPHM	366
Db	328	YRVDMTVPTAMTGGQDMLSNPBDVKMLLSEVNTLIYHKNIPWAHVDIFWGLDAPHM	387
Qy	367	YNEIHHMQQ 376	
Db	388	YNEIHHMQQ 397	

RESULT 9
 US-10-051-874-92
 Sequence 92, Application US/10051874
 Publication No. US2004000557A1
 GENERAL INFORMATION:
 APPLICANT: Padigaru, Muralidhara
 APPLICANT: Alsbrook II, John P
 APPLICANT: Colman, Steven D
 APPLICANT: Spytek, Kimberly A
 APPLICANT: Boldog, Ferenc
 APPLICANT: Vernet, Corine AM
 APPLICANT: Li, Li
 APPLICANT: Shenoy, Suresh G
 APPLICANT: Casman, Stacie J
 APPLICANT: Guo, Xiaojia Saaha
 APPLICANT: Edinger, Shlomit R
 APPLICANT: MacDougall, John R
 APPLICANT: Malyankar, Uriel M
 APPLICANT: Patturajan, Meera
 APPLICANT: Shinkets, Richard A
 APPLICANT: Pena, Carol EA
 APPLICANT: Tchernov, Velizar T
 APPLICANT: Millet, Isabelle
 APPLICANT: Miller, Charles E
 APPLICANT: Lepley, Denise M
 APPLICANT: Smithson, Glennda
 APPLICANT: Baumgartner, Jason C
 APPLICANT: Herrman, John L
 APPLICANT: Gorman, Linda
 APPLICANT: Mezes, Peter D
 APPLICANT: Kekuda, Ramesh
 APPLICANT: Taupier Jr, Raymond J
 APPLICANT: Gerlach, Valerie
 APPLICANT: Grosse, William M
 APPLICANT: Liu, Xiaohong
 APPLICANT: Ellerman, Karen
 APPLICANT: Rothenberg, Mark
 APPLICANT: Stone, David J
 APPLICANT: Burgess, Catherine E
 TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
 TITLE OF INVENTION: USING THE SAME
 FILE REFERENCE: 21402-245
 CURRENT APPLICATION NUMBER: US/10/051,874
 CURRENT FILING DATE: 2002-09-25
 PRIOR APPLICATION NUMBER: 60/268,595
 PRIOR FILING DATE: 2001-02-14
 PRIOR APPLICATION NUMBER: 60/325,306
 PRIOR FILING DATE: 2001-09-27
 PRIOR APPLICATION NUMBER: 60/262,587
 PRIOR FILING DATE: 2001-01-18
 PRIOR APPLICATION NUMBER: 60/272,409
 PRIOR FILING DATE: 2001-02-28
 PRIOR APPLICATION NUMBER: 60/262,454
 PRIOR FILING DATE: 2001-01-18
 PRIOR APPLICATION NUMBER: 60/276,777
 PRIOR FILING DATE: 2001-03-16
 PRIOR APPLICATION NUMBER: 60/291,672
 PRIOR FILING DATE: 2001-05-17
 PRIOR APPLICATION NUMBER: 60/330,336
 PRIOR FILING DATE: 2001-10-18
 PRIOR APPLICATION NUMBER: 60/265,530
 PRIOR FILING DATE: 2001-01-31
 PRIOR APPLICATION NUMBER: 60/261,376
 PRIOR FILING DATE: 2001-01-16
 NUMBER OF SEQ ID NOS: 269
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 92
 LENGTH: 399
 TYPE: PRT
 ORGANISM: Homo sapiens

US-10-051-874-92

Query Match 62.1%; Score 1289; DB 12; Length 399;
 Best Local Similarity 63.5%; Pred. No. 9.7e-129;
 Matches 235; Conservative 55; Mismatches 80; Indels 0; Gaps 0;

QY 7 AVDPAPMNISEIIHQGYPCBEYEVATEDGYILSVNRIPRGLVQPKTGRPVVLLQHG 66
 DB 28 AVDPETNMNVSEIISYWGFPSEYLVETEDGYILCLNIPRGRKNSDKGPKFVFLQHG 87

QY 67 LVGGASNWNLSNLPNNLSGLFADAGFDVWNGSRGNWRSKHKTLSDODEFWAFSYDEM 126
 DB 88 LLDSSNWNVTNLANSLSGLFADAGFDVWNGSRGNWRSKHKTLSDODEFWAFSYDEM 147

QY 127 ARFDLPAVINFILOKTOGEBKIYVYVGSQGTMTGFIAPSTMPPELAQKIKMYFALAPIATVK 186
 DB 148 AKYDLPASINFILOKTOGEBKIYVYVGSQGTMTGFIAPSTMPPELAQKIKMYFALAPIATVK 207

QY 187 HAKSPGTFKLLPDMWIKGLFGKKEFLYQTRFLRQLVYLCCGVILDDQICSNIMLLGGF 246
 DB 208 FCTSPMAKGLRPLDHLIKDLFGDKEFLPQSAFLKWLGHVTHVILKELCGLNLCFLLCGF 267

QY 247 NTNNMNSRASVYAAHTLAGTSVONILHWSQAVNSGELRAFDPWGSSEKTKNLEKNCQPTPVR 306
 DB 268 NERNLNSRDVYVYTHSPAGTSVQNMLHWSQAVKFKQAFDWSGSSAKNYFHYNQSYPT 327

QY 307 YRVDMVTPTAMWTGGODWLSNPEDVKMLLSEVTNLIYHKNIPWAHVDFIWLGLDAPHRM 366
 DB 328 YNVKMDLVPTAVWNGHDLADYVDVNYILLTQITNLVFNHESIPWEHLDFIWLGLDAPHRM 387

QY 367 YNEIHLMOQ 376
 DB 388 YNKIINLMRK 397

RESULT 10

US-10-051-874-93
 ; Sequence 93, Application US/10051874
 ; Publication No. US2004000557A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Padigaru, Muralidhara
 ; APPLICANT: Alsobrook II, John P
 ; APPLICANT: Colman, Steven D
 ; APPLICANT: Spyttek, Kimberly A
 ; APPLICANT: Boldog, Ferenc
 ; APPLICANT: Vernet, Corine AM
 ; APPLICANT: Li, Li
 ; APPLICANT: Shenoy, Suresh G
 ; APPLICANT: Casman, Stacie J
 ; APPLICANT: Guo, Xiaojia Sasha
 ; APPLICANT: Edinger, Shlomit R
 ; APPLICANT: MacDougall, John R
 ; APPLICANT: Malyankar, Uriel M
 ; APPLICANT: Patturajan, Meera
 ; APPLICANT: Shimkets, Richard A
 ; APPLICANT: Pena, Carol EA
 ; APPLICANT: Tchernev, Velizar T
 ; APPLICANT: Zerhusen, Bryan D
 ; APPLICANT: Millet, Isabelle
 ; APPLICANT: Miller, Charles E
 ; APPLICANT: Lepley, Denise M
 ; APPLICANT: Smithson, Glenda
 ; APPLICANT: Baumgarner, Jason C
 ; APPLICANT: Herrman, John L
 ; APPLICANT: Peyman, John A
 ; APPLICANT: Gorman, Linda
 ; APPLICANT: Mezes, Peter D
 ; APPLICANT: Kekuda, Rameesh
 ; APPLICANT: Taupier Jr, Raymond J
 ; APPLICANT: Gerlach, Valerie
 ; APPLICANT: Grosse, William M
 ; APPLICANT: Liu, Xiaohong
 ; APPLICANT: Ellerman, Karen

APPLICANT: Rothenberg, Mark
 APPLICANT: Stone, David J
 APPLICANT: Burgess, Catherine E
 TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
 TITLE OF INVENTION: USING THE SAME
 FILE REFERENCE: 21402-245
 CURRENT APPLICATION NUMBER: US/10/051,874
 CURRENT FILING DATE: 2002-09-25
 PRIOR APPLICATION NUMBER: 60/268,595
 PRIOR FILING DATE: 2001-02-14
 PRIOR APPLICATION NUMBER: 60/325,306
 PRIOR FILING DATE: 2001-09-27
 PRIOR APPLICATION NUMBER: 60/262,587
 PRIOR FILING DATE: 2001-01-18
 PRIOR APPLICATION NUMBER: 60/272,409
 PRIOR FILING DATE: 2001-02-28
 PRIOR APPLICATION NUMBER: 60/262,454
 PRIOR FILING DATE: 2001-01-18
 PRIOR APPLICATION NUMBER: 60/276,777
 PRIOR FILING DATE: 2001-03-16
 PRIOR APPLICATION NUMBER: 60/291,672
 PRIOR FILING DATE: 2001-05-17
 PRIOR APPLICATION NUMBER: 60/330,336
 PRIOR FILING DATE: 2001-10-18
 PRIOR APPLICATION NUMBER: 60/265,530
 PRIOR FILING DATE: 2001-01-31
 PRIOR APPLICATION NUMBER: 60/261,376
 PRIOR FILING DATE: 2001-01-16
 NUMBER OF SEQ ID NOS: 269
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 93
 LENGTH: 399
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-051-874-93

Query Match 62.1%; Score 1289; DB 12; Length 399;
 Best Local Similarity 63.5%; Pred. No. 9.7e-129;
 Matches 235; Conservative 55; Mismatches 80; Indels 0; Gaps 0;

QY 7 AVDPAPMNISEIIHQGYPCBEYEVATEDGYILSVNRIPRGLVQPKTGRPVVLLQHG 66
 DB 28 AVDPETNMNVSEIISYWGFPSEYLVETEDGYILCLNIPRGRKNSDKGPKFVFLQHG 87

QY 67 LVGGASNWNLSNLPNNLSGLFADAGFDVWNGSRGNWRSKHKTLSDODEFWAFSYDEM 126
 DB 88 LLDSSNWNVTNLANSLSGLFADAGFDVWNGSRGNWRSKHKTLSDODEFWAFSYDEM 147

QY 127 ARFDLPAVINFILOKTOGEBKIYVYVGSQGTMTGFIAPSTMPPELAQKIKMYFALAPIATVK 186
 DB 148 AKYDLPASINFILOKTOGEBKIYVYVGSQGTMTGFIAPSTMPPELAQKIKMYFALAPIATVK 207

QY 187 HAKSPGTFKLLPDMWIKGLFGKKEFLYQTRFLRQLVYLCCGVILDDQICSNIMLLGGF 246
 DB 208 FCTSPMAKGLRPLDHLIKDLFGDKEFLPQSAFLKWLGHVTHVILKELCGLNLCFLLCGF 267

QY 247 NTNNMNSRASVYAAHTLAGTSVONILHWSQAVNSGELRAFDPWGSSEKTKNLEKNCQPTPVR 306
 DB 268 NERNLNSRDVYVYTHSPAGTSVQNMLHWSQAVKFKQAFDWSGSSAKNYFHYNQSYPT 327

QY 307 YRVDMVTPTAMWTGGODWLSNPEDVKMLLSEVTNLIYHKNIPWAHVDFIWLGLDAPHRM 366
 DB 328 YNVKMDLVPTAVWNGHDLADYVDVNYILLTQITNLVFNHESIPWEHLDFIWLGLDAPHRM 387

QY 367 YNEIHLMOQ 376
 DB 388 YNKIINLMRK 397

RESULT 11
 US-10-051-874-94
 ; Sequence 94, Application US/10051874
 ; Publication No. US2004000557A1

GENERAL INFORMATION:

APPLICANT: Padigaru, Muralidhara
 APPLICANT: Alsobrook II, John P
 APPLICANT: Colman, Steven D
 APPLICANT: Spytek, Kimberly A
 APPLICANT: Boldog, Ferenc
 APPLICANT: Vernet, Corine AM
 APPLICANT: Li, Li
 APPLICANT: Shenoy, Suresh G
 APPLICANT: Casman, Stacie J
 APPLICANT: Guo, Xiaojia Sasha
 APPLICANT: Edinger, Shlomit R
 APPLICANT: MacDougall, John R
 APPLICANT: Malyankar, Uziel M
 APPLICANT: Patturajan, Meera
 APPLICANT: Shimkets, Richard A
 APPLICANT: Pena, Carol EA
 APPLICANT: Tchernev, Velizar T
 APPLICANT: Zethusen, Bryan D
 APPLICANT: Miliet, Isabelle
 APPLICANT: Miller, Charles E
 APPLICANT: Lepley, Denise M
 APPLICANT: Smithson, Glennnda
 APPLICANT: Baumgartner, Jason C
 APPLICANT: Herrman, John L
 APPLICANT: Peyman, John A
 APPLICANT: Gorman, Linda
 APPLICANT: Mezes, Peter D
 APPLICANT: Kekuda, Ramesh
 APPLICANT: Taupier Jr, Raymond J
 APPLICANT: Gerlach, Valerie
 APPLICANT: Grosse, William M
 APPLICANT: Liu, Xiaohong
 APPLICANT: Ellerman, Karen
 APPLICANT: Rothenberg, Mark
 APPLICANT: Stone, David J
 APPLICANT: Burgess, Catherine E
 TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
 FILE OF INVENTION: USING THE SAME
 FILE REFERENCE: 21402-245
 CURRENT APPLICATION NUMBER: US/10/051.874
 CURRENT FILING DATE: 2002-09-25
 PRIOR APPLICATION NUMBER: 60/268,595
 PRIOR FILING DATE: 2001-02-14
 PRIOR APPLICATION NUMBER: 60/325,306
 PRIOR FILING DATE: 2001-09-27
 PRIOR APPLICATION NUMBER: 60/262,587
 PRIOR FILING DATE: 2001-01-18
 PRIOR APPLICATION NUMBER: 60/272,409
 PRIOR FILING DATE: 2001-02-28
 PRIOR APPLICATION NUMBER: 60/262,454
 PRIOR FILING DATE: 2001-01-18
 PRIOR APPLICATION NUMBER: 60/276,777
 PRIOR FILING DATE: 2001-03-16
 PRIOR APPLICATION NUMBER: 60/291,672
 PRIOR FILING DATE: 2001-05-17
 PRIOR APPLICATION NUMBER: 60/330,336
 PRIOR FILING DATE: 2001-10-18
 PRIOR APPLICATION NUMBER: 60/265,530
 PRIOR FILING DATE: 2001-01-31
 PRIOR APPLICATION NUMBER: 60/261,376
 PRIOR FILING DATE: 2001-01-16
 NUMBER OF SEQ ID NOS: 269
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 94
 LENGTH: 399
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-051-874-94

Query Match 62.1%; Score 1289; DB 12; Length 399;
 Best Local Similarity 63.5%; Pred. No. 9.7e-129;
 Matches 235; Conservative 55; Mismatches 80; Indels 0; Gaps 0;

QY 7 AVDPFAFMNISEIIHQGYPCPEYEVATEDGYILSVNRIPRGLVQPKTGSPPVVLQHG 66
 DB 28 AVDPETNMVSEIIISYWGPFSEIYEVATEDGYILCLNRIPIHGRKNHSDKGPVVFVLOHG 87
 QY 67 LVGGASNIWISNLPNNISLGFILADAGFDVWNGSRGNWSRKHKTLSIDQDEFWAFSYDEM 126
 DB 88 LLADSSNWTNLANSLSLGFILADAGFDVWNGSRGNWSRKHKTLSVSDQDEFWAFSYDEM 147
 QY 127 ARFDLPVAVINFILQKTQGEKIYYVGYSGQTTMGFTAFSTMPBLAQKIMYFALAPIATVK 186
 DB 148 AKYDLPASINFILNKTQGEQVYVGHSGQTTGFTAFSQIPELAKRIKMFALGFVASVA 207
 QY 187 HAKSPGTCELLLPDMXIKLFGKKEFLYQTRFLRQLVYLCOQVILDOIICSNIMLLGGF 246
 DB 208 FCTSPMAKLGRLPDHLIKDLFGDKFLQSAFLKWLGVTHVCTHVLKELCGMLCFLCGF 267
 QY 247 NTNNMWSRASYAAHTLAGTSVQNIHWSQAVNGSELPAFDWGETKNLEKCNQPTVVR 306
 DB 268 NERNLMSRVDVYTTSTPAGTSVQNMHWSQAVKFKQAFDWSGSAKNIYHNGSYPT 327
 QY 307 YRVDMTVPMTAMWTCGQWLSPEDVKMLLSVTNLIYHKNIPWAHVDFTWGLDAPHRM 366
 DB 328 YNVKMLVPTAVASGHDMLADVDVNIILTOITNLVFHESIPWEHLDFTWGLDAPWRL 387
 QY 367 YNEIHLMOQ 376
 DB 388 YNKIINLMEK 397

RESULT 12
 US-10-042-431-41
 ; Sequence 41, Application US/10042431
 ; Publication NO. US20020182675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MCCARTHY, Sean A
 ; APPLICANT: BARNES, Thomas M
 ; APPLICANT: FRASER, Christopher C
 ; APPLICANT: SHARP, John D
 ; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
 ; FILE OF INVENTION: PREVENTIVE, THERAPEUTIC, AND OTHER USES
 ; FILE REFERENCE: 10147-6U2
 ; CURRENT APPLICATION NUMBER: US/10/042,431
 ; CURRENT FILING DATE: 2001-10-25
 ; PRIOR APPLICATION NUMBER: US 09/333,159
 ; PRIOR FILING DATE: 1999-06-14
 ; PRIOR APPLICATION NUMBER: US 09/578,063
 ; PRIOR FILING DATE: 2000-05-24
 ; NUMBER OF SEQ ID NOS: 79
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 41
 ; LENGTH: 399
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-042-431-41

Query Match 62.1%; Score 1289; DB 14; Length 399;
 Best Local Similarity 63.5%; Pred. No. 9.7e-129;
 Matches 235; Conservative 55; Mismatches 80; Indels 0; Gaps 0;

QY 7 AVDPFAFMNISEIIHQGYPCPEYEVATEDGYILSVNRIPRGLVQPKTGSPPVVLQHG 66
 DB 28 AVDPETNMVSEIIISYWGPFSEIYEVATEDGYILCLNRIPIHGRKNHSDKGPVVFVLOHG 87
 QY 67 LVGGASNIWISNLPNNISLGFILADAGFDVWNGSRGNWSRKHKTLSIDQDEFWAFSYDEM 126
 DB 88 LLADSSNWTNLANSLSLGFILADAGFDVWNGSRGNWSRKHKTLSVSDQDEFWAFSYDEM 147
 QY 127 ARFDLPVAVINFILQKTQGEKIYYVGYSGQTTMGFTAFSTMPBLAQKIMYFALAPIATVK 186
 DB 148 AKYDLPASINFILNKTQGEQVYVGHSGQTTGFTAFSQIPELAKRIKMFALGFVASVA 207
 QY 187 HAKSPGTCELLLPDMXIKLFGKKEFLYQTRFLRQLVYLCOQVILDOIICSNIMLLGGF 246

```

Db 208 FCTSPMAKGLRDLKDFGDFLQSAFLKLGTHVTHVILKELCGNLCFLGCF 267
Qy 247 NTNNMMSRASVYAAHTLAGTSVONILHWSQAVNSGELAPDWSGSKNLEKCNQPTVR 306
Db 268 NERNLNMSRDVVTTHSPAGTSVQNNMLHWSQAVKQFQAFDWSGSAKKNYFHNQSYPT 327
Qy 307 YRVDMTVPMTAMTGGDMLSNPDEYKMLISEVTNLIYHKNIPENAHVDFIWLGLDAPHRM 366
Db 328 YNVKMDLVPVAVWSGGHDLADVVDVNNILITQITNLVPHESIPWEHDLFIWLGLDAPWRL 387
Qy 367 YNEIHLMOQ 376
Db 388 YNKIINLMRK 397

```

```

RESULT 13
US-10-003-302-4
; Sequence 4, Application US/10003302
; Publication No. US20020142435A1
; GENERAL INFORMATION:
; APPLICANT: MERKULOV, Gennady et al.
; TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001186DIV
; CURRENT APPLICATION NUMBER: US/10/003,302
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-003-302-4

```

```

Query Match 61.8%; Score 1282; DB 14; Length 392;
Best Local Similarity 63.8%; Pred. No. 5.3e-128;
Matches 234; Conservative 53; Mismatches 80; Indels 0; Gaps 0;

Qy 7 AVDPAPFNISEIIHQGYPCPEYEVATEDGYILSVNRIPRGLVQPKTGSRRPVLLQHG 66
Db 26 AVDPETNMYSEIISYWGFPSEBYLVETEDGYILCLNRIPIHGRKNHSDKGPVWFLQHG 85
Qy 67 LVGASNWSNLNPNLSGLFADAGFDVWNGNSRGNWSRKHKTLSIDQDEFWAFSYDEM 126
Db 86 LLADSSNWTNLANSLSGLFADAGFDVWNGNSRGNWTSEKHTLSVQDEFWAFSYDEM 145
Qy 127 ARFDLPVAVINFILOKTOGQEKIYVYVGSQGTMTGFIAPSTMPDLAKIKMYFALAPATVK 186
Db 146 AKYDLPASINFILOKTOGQEKIYVYVGSQGTMTGFIAPSTMPDLAKIKMYFALGPVAVSA 205
Qy 187 HAKSPGTFKLLPDMMIKGLFGKKEFLYQTRFLRQLVYLCOGVIDLQICSNIMLLGCF 246
Db 206 FCTSPMAKGLRDLKDFGDFLQSAFLKLGTHVTHVILKELCGNLCFLGCF 265
Qy 247 NTNNMMSRASVYAAHTLAGTSVONILHWSQAVNSGELAPDWSGSKNLEKCNQPTVR 306
Db 266 NERNLNMSRDVVTTHSPAGTSVQNNMLHWSQAVKQFQAFDWSGSAKKNYFHNQSYPT 325
Qy 307 YRVDMTVPMTAMTGGDMLSNPDEYKMLISEVTNLIYHKNIPENAHVDFIWLGLDAPHRM 366
Db 326 YNVKMDLVPVAVWSGGHDLADVVDVNNILITQITNLVPHESIPWEHDLFIWLGLDAPWRL 385
Qy 367 YNEIHL 373
Db 386 YNKIINL 392

```

```

RESULT 14
US-09-759-130B-420
; Sequence 420, Application US/09759130B
; Publication No. US2003002279A1

```

```

; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A
; APPLICANT: Fraser, Christopher C
; APPLICANT: Sharp, John D
; APPLICANT: Barries, Thomas S
; APPLICANT: Kirt, Susan J
; APPLICANT: Mackay, Charles R
; APPLICANT: Myers, Paul S
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; TITLE OF INVENTION: USES
; FILE REFERENCE: MP100-535OMNIM
; CURRENT APPLICATION NUMBER: US/09/759,130B
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/602,871
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/420,707
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 420
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-759-130B-420

Query Match 56.1%; Score 1165; DB 11; Length 221;
Best Local Similarity 100.0%; Pred. No. 7.1e-116;
Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VHMPTKAVDPEAFNISEIIHQGYPCPEYEVATEDGYILSVNRIPRGLVQPKTGSRRPV 60
Db 1 VHMPTKAVDPEAFNISEIIHQGYPCPEYEVATEDGYILSVNRIPRGLVQPKTGSRRPV 60
Qy 61 VLLQHLVGGASNWSNLNPNLSGLFADAGFDVWNGNSRGNWSRKHKTLSIDQDEFWA 120
Db 61 VLLQHLVGGASNWSNLNPNLSGLFADAGFDVWNGNSRGNWSRKHKTLSIDQDEFWA 120
Qy 121 PSYDEMARFDLPVAVINFILOKTOGQEKIYVYVGSQGTMTGFIAPSTMPDLAKIKMYFALA 180
Db 121 PSYDEMARFDLPVAVINFILOKTOGQEKIYVYVGSQGTMTGFIAPSTMPDLAKIKMYFALA 180
Qy 181 PIATVGHAKSPGTFKLLPDMMIKGLFGKKEFLYQTRFLRQ 221
Db 181 PIATVGHAKSPGTFKLLPDMMIKGLFGKKEFLYQTRFLRQ 221

```

```

RESULT 15
US-10-042-431-50
; Sequence 50, Application US/10042431
; Publication No. US20020182675A1
; GENERAL INFORMATION:

```

; APPLICANT: MCCARTHY, Sean A
 ; APPLICANT: BARNES, Thomas M
 ; APPLICANT: FRASER, Christopher C
 ; APPLICANT: SHARP, John D
 ; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
 ; TITLE OF INVENTION: PREVENTIVE, THERAPEUTIC, AND OTHER USES
 ; FILE REFERENCE: 10147-6U2
 ; CURRENT APPLICATION NUMBER: US/10/042,431
 ; CURRENT FILING DATE: 2001-10-25
 ; PRIOR APPLICATION NUMBER: US 09/333,159
 ; PRIOR FILING DATE: 1999-06-14
 ; PRIOR APPLICATION NUMBER: US 09/578,063
 ; PRIOR FILING DATE: 2000-05-24
 ; NUMBER OF SEQ ID NOS: 79
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 50
 ; LENGTH: 221
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-042-431-50

Query Match 56.1%; Score 1165; DB 14; Length 221;
 Best Local Similarity 100.0%; Pred. No. 7.1e-116;
 Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 Qy 1 VHMPTKAVDPEAFNISEIIHQGYPCPEYEYEVATEDGYILSVNRIPRGLVQPKKTSRPPV 60
 Db 1 VHMPTKAVDPEAFNISEIIHQGYPCPEYEYEVATEDGYILSVNRIPRGLVQPKKTSRPPV 60
 Qy 61 VLLQHLVGGASNWSNLPNNLSGLFILADAGFDVWMCNSRGNWSRKHKTLISIDQDEFWA 120
 Db 61 VLLQHLVGGASNWSNLPNNLSGLFILADAGFDVWMCNSRGNWSRKHKTLISIDQDEFWA 120
 Qy 121 PSYDEMARFDLPVAVINFILOKTQGEKIYYGYSGITMGFIAPSTWPELAQKIKMYFALA 180
 Db 121 PSYDEMARFDLPVAVINFILOKTQGEKIYYGYSGITMGFIAPSTWPELAQKIKMYFALA 180
 Qy 181 PIATVKHAKSPGTKFLLLPDMWIKGLFGKKEFLYQTRFLRQ 221
 Db 181 PIATVKHAKSPGTKFLLLPDMWIKGLFGKKEFLYQTRFLRQ 221

Search completed: February 19, 2004, 14:34:26
 Job time : 28.3838 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2004, 14:17:49 ; Search time 13.4317 Seconds
(without alignments)
2792.326 Million cell updates/sec

Title: US-10-042-431-49

Perfect score: 2076
Sequence: 1 VHMPTKAVDPEAFNMISEII.....IHLMOQBETNLSQRCBAVL 390

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_76:*
1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1289	62.1	399	2 S41408	lysosomal acid lip
2	1289	62.1	399	2 G01416	lysosomal acid lip
3	1161	55.9	395	1 L1RTT	triacylglycerol li
4	1125	54.2	397	1 J24017	triacylglycerol li
5	1109	53.4	398	2 S07145	triacylglycerol li
6	747	36.0	411	2 T22290	hypothetical prote
7	744.5	35.9	403	2 T33198	hypothetical prote
8	727	35.0	405	2 T22675	hypothetical prote
9	699	33.7	405	2 H88930	protein R11G11.14
10	691.5	33.3	411	2 G89074	protein K04A8.5 [i
11	651.5	31.4	426	2 T30480	hypothetical prote
12	574	27.6	559	2 J70949	egg-specific prote
13	491	23.7	443	2 T39540	triglyceride lipas
14	478.5	23.0	1585	2 T31611	hypothetical prote
15	475	22.9	467	2 T41053	triglyceride lipas
16	445	21.4	344	2 E84526	probable lysosomal
17	422.5	20.4	548	2 S37969	probable triacylgl
18	403	19.4	460	2 T39443	probable triacylgl
19	397.5	19.1	413	2 T33170	probable triacylgl
20	294	14.2	431	2 D86318	protein F5H18.6 [
21	282	13.6	538	2 S64842	probable membrane
22	271.5	13.1	573	2 S64754	protein lipase F2p
23	232.5	11.2	509	2 G96766	lipase, pregastric
24	149	7.2	62	2 S59904	probable esterase/
25	143	6.9	336	2 P83425	conserved hypothet
26	113	5.4	412	2 B72391	triacylglycerol li
27	112.5	5.4	40	2 S29539	hypothetical prote
28	111.5	5.4	987	2 A64474	lysophospholipase
29	107	5.2	281	2 A19112	

30	105.5	5.1	304	2 A99905	hypothetical prote
31	105	5.1	430	2 T23001	hypothetical prote
32	103.5	5.0	798	2 F96714	probable protease
33	100.5	4.8	929	2 JH0262	Outer protein - Eme
34	99.5	4.8	286	2 D75217	probable 2-acetyl-
35	99	4.8	318	2 AE2968	ABC transporter, m
36	99	4.8	318	2 F98314	probable permease
37	99	4.8	460	2 G96764	unknown protein P2
38	98.5	4.7	436	2 JC5021	platelet-activatin
39	97.5	4.7	322	2 B70957	probable spha prot
40	97.5	4.7	456	2 E63991	hypothetical prote
41	96.5	4.6	287	1 E64053	hypothetical prote
42	96	4.6	809	2 S67153	probable membrane
43	96	4.6	934	2 B82364	DNA polymerase I V
44	95	4.6	596	2 AD1876	hypothetical prote
45	94.5	4.6	1275	2 I38588	reverse transcript

ALIGNMENTS

RESULT 1

S41408
lysosomal acid lipase (EC 3.1.1.1.-) / sterol esterase (EC 3.1.1.13) precursor - human
C/Species: Homo sapiens (man)
C/Date: 13-Jan-1995 #sequence revision 13-Jan-1995 #text_change 18-Jun-1999
C/Accession: S41408; A39315; S47187
R/Aneis, D.; Merkel, M.; Eckerskorn, C.; Greten, H.
Eur. J. Biochem. 219, 905-914, 1994
A/Title: Purification, characterization and molecular cloning of human hepatic lysosomal
A/Reference number: S41408; MUID:94155897; PMID:8112342
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-399 <AME>
A/Cross-references: EMBL:X76488; NID:G4344305; PIDN:CAA54026.1; PID:G4344306
R/Anderson, R.A.; Sando, G.N.
J. Biol. Chem. 266, 22479-22484, 1991
A/Title: Cloning and expression of cDNA encoding human lysosomal acid lipase/cholesteryl
A/Reference number: A39315; MUID:92042192; PMID:1718995
A/Accession: A39315
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-15, 'P', 17-399 <AND>
A/Cross-references: GB:M74775; NID:G187151; PIDN:AAA59519.1; PID:G187152
R/Du, H.; Gregory, G.A.
submitted to the EMBL Data Library, April 1994
A/Description: Structural conservation of putative functional motifs between mouse and h
A/Reference number: S47187
A/Accession: S47187
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-22, 'R', 24-399 <DUH>
A/Cross-references: EMBL:231690; NID:G506430; PIDN:CAA83495.1; PID:G506431
C/Genetics:
A/Gene: GDB:LIPA
A/Cross-references: GDB:120153; OMIM:278000
A/Map position: 10q24-10q25
C/Superfamily: triacylglycerol lipase, lingual
C/Keywords: carboxylic ester hydrolase; glycoprotein

Query Match	62.1%	Score 1289;	DB 2;	Length 399;
Best Local Similarity	63.5%	Fred. No. 8.9e-107;		
Matches 235;	Conservative 55;	Mismatches 80;	Indels 0;	Gaps 0;
Qy	7	AVDPEAFNMISEIIHQGYPCEEYEVATEDGYILSVNRIIPRGLVQPKTKGSPVVLQHG	66	
Db	28	AVDPEATNMVSVRIISVWGPFPSEYLVEEDGYILCLNRIHPHGRKNHSDKGPVVLQHG	87	
Qy	67	LVGGSANWISNLPNNISGLFILADAGFDVWNGSRGNWRSRKHTLSIQODEFWAFSYDEM	126	
Db	88	LLADSSNWTNLSGLFILADAGFDVWNGSRGNWRSRKHTLSVQODEFWAFSYDEM	147	

QY 127 ARFDLPAVINFILOKTQGEKIYVYVGSQGTMTGFIAPFTWPELAQKIMYFALAPATYK 186
 Db 148 AKYDLPASINFILOKTQGEKIYVYVGSQGTMTGFIAPFTWPELAQKIMYFALAPATYK 207
 QY 187 HAKSPGKFLLLPDMNKGKFGKKEFLYQTRFLRQVILVCGVILDOICSNIMLLGQF 246
 Db 208 FCTSPMAKGLPDLHLKIDFGKKEFLYQTRFLRQVILVCGVILDOICSNIMLLGQF 267
 QY 247 NTNNMNSRVSVAHAHTLAGTSVQNLHWSQAVNSGELRAFDWGSSTKLEKCNQPTPYR 306
 Db 268 NERNLNSRDVVTTHSPAGTSVQNLHWSQAVNSGELRAFDWGSSTKLEKCNQPTPYR 327
 QY 307 YRVDMVTPTAMTGGQDNLSPEDVKMLLSEVTLNLYHKNIPPEWAVDFPIGLDAPHRM 366
 Db 328 YNVKMDLVPVAVWSGGHDWLDVYDVNILLTQTNLVFHEIPSEWHELDPIGLDAPHRM 387
 QY 367 YNEIHLMOQ 376
 Db 388 YNKIINLMRK 397

RESULT 2

G01416
 lysosomal acid lipase - human
 C:Species: Homo sapiens (man)
 C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 18-Jun-1999
 C:Accession: G01416
 R:Du, H.

submitted to the EMBL Data Library, April 1994

A:Reference number: G06919
 A:Accession: G01416
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-399 <DUX>
 A:Cross-references: EMBL:U08464; NID:G505052; PIDN:AA60328.1; PID:G505053
 C:Superfamily: triacylglycerol lipase, lingual
 C:Keywords: Glycoprotein

Query Match 62.1%; Score 1289; DB 2; Length 399;
 Best Local Similarity 63.5%; Pred. No. 8.9e-107;
 Matches 235; Conservative 55; Mismatches 80; Indels 0; Gaps 0;
 QY 7 AVDEAFWNISIIHQGYPCREYEVATEDGYILSVNRIPRGVQPKTGRSPVLLQHG 66
 Db 28 AVDETNMNSIISYGFPSSEYLVETEDGYILCLNRPHGKNSDKGPKPFWVLQHG 87
 QY 67 LVGASNWIINLPNNSLGFILADAGFDVWNGSRGNAMSRKHKTLSIDODEFWAFSYDEM 126
 Db 88 LLADSSNWVTLNANSLGFILADAGFDVWNGSRGNAMSRKHKTLSVQDFWAFSYDEM 147
 QY 127 ARFDLPAVINFILOKTQGEKIYVYVGSQGTMTGFIAPFTWPELAQKIMYFALAPATYK 186
 Db 148 AKYDLPASINFILOKTQGEKIYVYVGSQGTMTGFIAPFTWPELAQKIMYFALAPATYK 207
 QY 187 HAKSPGKFLLLPDMNKGKFGKKEFLYQTRFLRQVILVCGVILDOICSNIMLLGQF 246
 Db 208 FCTSPMAKGLPDLHLKIDFGKKEFLYQTRFLRQVILVCGVILDOICSNIMLLGQF 267
 QY 247 NTNNMNSRVSVAHAHTLAGTSVQNLHWSQAVNSGELRAFDWGSSTKLEKCNQPTPYR 306
 Db 268 NERNLNSRDVVTTHSPAGTSVQNLHWSQAVNSGELRAFDWGSSTKLEKCNQPTPYR 327
 QY 307 YRVDMVTPTAMTGGQDNLSPEDVKMLLSEVTLNLYHKNIPPEWAVDFPIGLDAPHRM 366
 Db 328 YNVKMDLVPVAVWSGGHDWLDVYDVNILLTQTNLVFHEIPSEWHELDPIGLDAPHRM 387

RESULT 3

LIRY

triacylglycerol lipase (EC 3.1.1.3) precursor, lingual - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 11-Jun-1999
 C:Accession: A23045
 R:Doeherty, A.J.P.; Bodmer, M.W.; Angal, S.; Verger, R.; Riviere, C.; Lowe, P.A.; Lyons, Nucleic Acids Res. 13, 1891-1903, 1985
 A:Title: Molecular cloning and nucleotide sequence of rat lingual lipase cDNA.
 A:Reference number: A23045; MUID:85215587; PMID:3839077
 A:Accession: A23045
 A:Molecule type: mRNA
 A:Residues: 1-395 <DOC>
 A:Cross-references: GB:X02309; NID:G56595; PIDN:CAA26179.1; PID:G56596
 A:Experimental source: serain Sprague-Dawley
 A:Note: The partial sequence of the mature protein from a different, unspecified strain
 C:Comment: This acid-stable lipase is secreted by the serous (von Ebner's) glands at the
 C:Superfamily: triacylglycerol lipase, lingual
 C:Keywords: carboxylic ester hydrolase; glycoprotein; lipid digestion; saliva; serous gl
 P:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>
 F:33,68,98,184,270/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 55.9%; Score 1161; DB 1; Length 395;
 Best Local Similarity 57.7%; Pred. No. 2.2e-95;
 Matches 213; Conservative 63; Mismatches 93; Indels 0; Gaps 0;

QY 9 DPBAPNISEIIHQGYPCREYEVATEDGYILSVNRIPRGVQPKTGRSPVLLQHG 68
 Db 27 NPBANMISQITWYGPCEYEVATEDGYILGYRIPHGKNSENIGKRPVYLQGLI 86
 QY 69 CGASNWIINLPNNSLGFILADAGFDVWNGSRGNAMSRKHKTLSIDODEFWAFSYDEM 128
 Db 87 ASATNWIINLPNNSLGFILADAGFDVWNGSRGNAMSRKHKTLSIDODEFWAFSYDEM 146
 QY 129 FDLPAVINFILOKTQGEKIYVYVGSQGTMTGFIAPFTWPELAQKIMYFALAPATYK 188
 Db 147 YDLPATNFILOKTQGEKIYVYVGSQGTMTGFIAPFTWPELAQKIMYFALAPATYK 206
 QY 189 KSPGKFLLLPDMNKGKFGKKEFLYQTRFLRQVILVCGVILDOICSNIMLLGQFNT 248
 Db 207 QSPKKSIFITFLFKLMFGKMFPHYTFDDLTGTEVCSREVLNLLCSNTLFIQGFDK 266
 QY 249 NNMNSRVSVAHAHTLAGTSVQNLHWSQAVNSGELRAFDWGSSTKLEKCNQPTPYR 308
 Db 267 KNLNNSRDVVTTHSPAGTSVQNLHWSQAVNSGELRAFDWGSSTKLEKCNQPTPYR 326
 QY 309 VRDMVTPTAMTGGQDNLSPEDVKMLLSEVTLNLYHKNIPPEWAVDFPIGLDAPHRM 368
 Db 327 VSAMTVPVAVWSGGHDWLDVYDVNILLTQTNLVFHEIPSEWHELDPIGLDAPHRM 386
 QY 369 EIIHLMQOE 377
 Db 387 EIMSNMAED 395

RESULT 4

JC4017
 triacylglycerol lipase (EC 3.1.1.3) PGF precursor - bovine
 N:Alternate names: pregastric esterase
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: JC4017; S64678
 R:Timmermans, M.Y.J.; Teuchy, H.; Kupers, L.P.M.
 Gene 147, 259-262, 1994
 A:Title: The cDNA sequence encoding bovine pregastric esterase.
 A:Reference number: JC4017; MUID:95011625; PMID:7926811
 A:Accession: JC4017
 A:Molecule type: mRNA
 A:Residues: 1-397 <TIM>
 A:Cross-references: GB:L26319; NID:G600756; PIDN:AAA57037.1; PID:G600757
 A:Experimental source: tongue
 R:Timmermans, M.Y.J.; Reekmans, G.; Teuchy, H.J.H.; Kupers, L.P.M.
 Biochem. J. 314, 931-936, 1996
 A:Title: Inhibition studies on calf pregastric esterase: the enzyme has no functional ti

A;Residues: 20-45 <BER>
C;Superfamily: triacylglycerol lipase, lingual

QY

5QGTMGF-
||| : |

Db 143 WDEMAQYDVPAMVDHVLMTAQENLYYMGHSGQGLTLMFTHLAKDTGDSFAKKIKRYFALA 202
 Qy 181 PIATVKHAKSPGPKFLLPDMWIKG---LFGKKEFLYQTRFLRQLVYICGQVILD-QIC 236
 Db 203 FIGAVRNKIGFLSYFAHKFSPEFDGMYELFGSKDFLFDNNWITKMAAKDICGASEKEAEIC 262
 Qy 237 SNTMLLGGFNTNNMNSRASVYAAHTLACTSVQNTILHWSQAVNSGELRAFQWGSSTKQL 296
 Db 263 DNELFLIGGPESQWNASRTAIYSSQDPAGTSTQNVHVMQVNRGRVPAFDWGGKI-NK 321
 Qy 297 EKNQPTPVRY---RVKDMTPVPAWMTGGQDWLSNPEDV-KMLLSEVTNLIYHKNI--PE 350
 Db 322 KKGQDTPPEYDFGAIGTKI--HLYWSDDMDLGDPTDHDHFLKELNPAVIAENNLKS 379
 Qy 351 WAHVDFIWLGLDAPHRMYNEI 371
 Db 380 YNHLDFTWGLNATFIYDKAI 400

RESULT 7
 T33198
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000
 C:Accession: T33198
 R:Wu, X.
 submitted to the EMBL Data Library, May 1998
 A:Description: The sequence of C. elegans cosmid ZK6.
 A:Reference number: Z21301
 A:Accession: T33198
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-403 <WUX>
 A:Cross-references: EMBL:AF067942; PIDN:AAC17694.1; GSPDB:GN00023; CESP:ZK6.7
 A:Experimental source: strain Bristol N2; clone ZK6
 C:Genetics:
 A:Gene: CESP:ZK6.7
 A:Map position: 5
 A:Introns: 29/3; 63/1; 219/3; 319/3; 365/3
 C:Superfamily: triacylglycerol lipase, lingual

Query Match 35.9%; Score 744.5; DB 2; Length 403;
 Best Local Similarity 42.3%; Pred. No. 2.8e-58;
 Matches 162; Conservative 62; Mismatches 132; Indels 27; Gaps 10;

Qy 9 DPEAFNMISEIIHQGYPCSEYEVATEDGYLSVNRIPRGLVQPKT-----GSRPVVL 62
 Db 22 DPELHMTTPQIIIRWGYPAMIYVATDDGYILEHRIIPG-----KTNVTWPNKGPVVF 76
 Qy 63 LQHGLVGGASNIWISNLPNSLGFILADAGFDVWNGSRGNWGRKKTLSIDQDEFWAPS 122
 Db 77 MQHGLLCASDWWVNLDPQAGFLFADAGFDVWLNWRGNTYSMKHKDLKPSHSAFDWS 136
 Qy 123 YDEMARFDLPVAVINFIQKQGEKIYVYGSQGTWGTAFSTMP-ELAKIKRYF 181
 Db 137 WDEMATYDLNAMINHLVLEVTGQDSVYVYMGHSQGLTMTFSLSKDQDGSFAKKIKKFPALAP 196
 Qy 182 IATVKHAKSPGPKFLLPDMWIKG---LFGKKEFLYQTRFLRQLVYICGQVILD-QICS 237
 Db 197 IGSVKHILKGLSPFANYFESLEFDGWFDFGAGFLPNWAMKLAADICGLKVEADLDC 256
 Qy 238 NIMLLGGFNTNNMNSRASVYAAHTLACTSVQNTILHWSQAVNSGELRAFQWGSSTKYLE 297
 Db 257 NVLFELIAGPESQWNASRTAIYSSQDPAGTSTQNVHVMQVNRGRVPAFDWGGTKI-NKK 315
 Qy 298 KKNQPTPVRYVRDWT-----VPTAMWTGGQDWLSNPEDV-KMLLSEVTNLIYHKNI--IPE 350
 Db 316 KYGANPEY---DFTAIKGYLYIYWSDDADWLADPTDVPDYLTLRLNPAVIAQNHLPD 372
 Qy 351 WAHVDFIWLGLDAPHRMYNEI 373
 Db 373 YNHLDFTWGLRAPDPIYRPAIKL 395

RESULT 8
 T22675
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
 C:Accession: T22675
 R:Percy, C.; Lloyd, C.
 submitted to the EMBL Data Library, September 1996
 A:Reference number: Z19598
 A:Accession: T22675
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-405 <WIL>
 A:Cross-references: EMBL:Z79696; PIDN:CAB01973.1; GSPDB:GN00023; CESP:F54F3.3
 A:Experimental source: clone F54F3
 C:Genetics:
 A:Gene: CESP:F54F3.3
 A:Map position: 5
 A:Introns: 31/3; 65/1; 274/3; 366/3
 C:Superfamily: triacylglycerol lipase, lingual

Query Match 35.0%; Score 727; DB 2; Length 405;
 Best Local Similarity 40.3%; Pred. No. 1e-56;
 Matches 157; Conservative 65; Mismatches 140; Indels 28; Gaps 10;

Qy 9 DPEAFNMISEIIHQGYPCSEYEVATEDGYLSVNRIPRGLVQPKT-----GSRPVVL 62
 Db 24 DPEKMTTPQIIIRWGYPAMIYVATDDGYILEHRIIPG-----KTNVTWPNKGPVVF 78
 Qy 63 LQHGLVGGASNIWISNLPNSLGFILADAGFDVWNGSRGNWGRKKTLSIDQDEFWAPS 122
 Db 79 MQHGLCSSNWVNLPTESAFLFADAGDYDVLGNFRGNTYSMKHKDLKPSHSAFDWS 138
 Qy 123 YDEMARFDLPVAVINFIQKQGEKIYVYGSQGT-TMGFIATSTMP-ELAKIKRYF 177
 Db 139 WDEMQQYDLPAIEKALEVTGQDSLYYIGHSQGLTMT-----FSRLSEDKVGMGNKIKKFF 194
 Qy 178 ALAPIATVKHAKSPGPKFLLPDMWIKG---LFGKKEFLYQTRFLRQLVYICGQVILD 233
 Db 195 ALAPVGSVKHILKGLSPFANYFESLEFDGWFDFGAGFLPNWAMKLAADICGLKVEA 254
 Qy 234 QICSNIMLLGGFNTNNMNSRASVYAAHTLACTSVQNTILHWSQAVNSGELRAFQWGSST 293
 Db 255 GVCDDVWFLIAGPESQWNASRTAIYSSQDPAGTSTQNVHVMQVNRGRVPAFDWGGTKI-EK 313
 Qy 294 KMLEKKNQPTPVRYVRDWTMTPTAMWTGGQDWLSNPEDV-KMLLSEV--TNLIYHKNIPE 350
 Db 314 GNKKHYGQANVPAYDFTTVNRFPVLYMGSDWLADPTDVTDFLLTLNLPSTVVQNNKLID 373
 Qy 351 WAHVDFIWLGLDAPHRMYNEI 380
 Db 374 YNHLDFIWLGLRAPDPIYRPAIKL 403

RESULT 9
 H88930
 C:Species: Caenorhabditis elegans
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
 C:Accession: H88930
 R:Anonymous, The C. elegans Sequencing Consortium.
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
 A:Reference number: A75000; MUID:99069613; PMID:9851916
 A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/projects/C_elegans/ and
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
 A:Accession: H88930
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-405 <STO>
 A:Cross-references: GB:chr V; PIDN:AAC69088.1; PID:g2384863; GSPDB:GN00023; CESP:R11G11.14
 A:Note: Similar to lipase; R11G11.14

C:Genetics:
A:Gene: R11G11.14
A:Map position: 5
C:Superfamily: triacylglycerol lipase, lingual

Query Match 33.7%; Score 699; DB 2; Length 405;
Best Local Similarity 39.5%; Pred. No. 3.1e-54;
Matches 158; Conservative 69; Mismatches 137; Indels 36; Gaps 13;

QY 1 VHMPTKAVDPBPAFMNISEIIHQGYPCPEYEVATEDGYLSVNRIPRGLVQPKKT----55
DB 15 VHVHAKSDPELHMTTPQIIERWGPAMYISVTDDGVILELHRIPHG-----KNTWTP 69
QY 56 -GSRPVLLQGLVGAGNWTISNPNNSLGFILADAGFDVVMGNSRGNAMSRKHKTLISID 114
DB 70 NGKQVWFMOHGLLCASDWTMLPEQSAFIFADAGFDVLMGNRGNTYSMKHKLKAS 129
QY 115 QDEWAFSDYDEMARFDLPAVINFILOKTQGEKIYVYVSGQT-TMGFIAPSTWPE----L 169
DB 130 HSDFEWSEWDEWATYDLPAMDVKLEVTGQESLYMGHSQGLTWM-----FSLHSDKDGIF 185
QY 170 AQTKIMYFALAPIATVGHAKSPGKFLLLPDMIMK---LFGKKEFLYQTRFLRQLVIYL 226
DB 186 AKKIKKFPALAPVGSVKDIKGLFFFAHYFSLFEDGWFDFGAGEFLPNWAMKLAADI 245
QY 227 CGQV-ILDOIICSNIMLLGGFTNNMNSRASVVAHTLAGTSVQNLHWSQAVNSGELR 285
DB 246 CGGLKIESDLCDNVCFIAGPESQDNSTRFVYASHDPAATATONIVHWIQMVYRHGGVP 305
QY 286 AFDGSETKMLEKCN-QTPVRYVRVDMT-----VPTAMWTGGQDWLGNPDEV-KMLSEV 339
DB 306 AYDMSK-ENKNYVFNQANPEY---DFTAKGQIVLYNSDADWLADKDTITNLLTRL 361
QY 340 TNLHYKYN--IPEWAVDFIWLGDAPHMYNEIHLMOQE 377
DB 362 NPATIAQNNYFTDYNHFDVFLRAPNDIYLPVICTKD 401

RESULT 10
G89074
protein K04A8.5 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
A:Accession: G89074
R:anonymus, The C. elegans Sequencing Consortium.
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A:Reference number: A75000; MUID:99069613; PMID:981916
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/projects/C.ele
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: G89074
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-411 <STO>
A:Cross-references: GB:chr_V; PIDN:AAC48051.1; PID:g1658354; GSPDB:GN00023; CESP:K04A8.5
C:Genetics:
A:Gene: K04A8.5
A:Map position: 5
C:Superfamily: triacylglycerol lipase, lingual

Query Match 33.3%; Score 691.5; DB 2; Length 411;
Best Local Similarity 38.6%; Pred. No. 1.5e-53;
Matches 150; Conservative 77; Mismatches 143; Indels 19; Gaps 9;

QY 5 TKAVDPEAFPMNISEIIHQGYPCPEYEVATEDGYLSVNRIPRGLVQPKKT---GSRPVV 61
DB 25 SKSDLEFLYLTPELISWGSVBIYNTTKDGFILHLRIPYGREVPTSSDVNNSPVI 84
QY 62 LLQGLVGAGNWTISNPNNSLGFILADAGFDVVMGNSRGNAMSRKHKTLISIDQ-EFWA 120
DB 85 FLOHGLCSSFDVANSFHQAGFADAGFDVLMGNRGNTYSRKHVSINLPKDPKFD 144
QY 121 PSYDEMARFDLPAVINFILOKTQGEKIYVYVSGQT-TMGFIAPSTWPELAQKIMYFALA 180

Db 145 WSWDQISEYDLFAMIGKALEISQSSLYTGSGLTMTFAKISTDPKFSRKIKYFALA 204
QY 181 PIATVGHAKSPGKFLLLPDMIMKGL-----FGKKEFLYQTRFLRQLVIYLCQV-ILD 233
Db 205 PIGSIKHAHG---VFLFLGRHFGKDYEVYKKGSDDELFGSSLLPKKIKYKTCGLFDLLE 261
QY 234 QICSNIMLLGGFTNNMNSRASVVAHTLAGTSVQNLHWSQAVNSGELRAFDPWGSSET 293
Db 262 EFCSDITLLFIGTANENWQTRIPVYLAHTPAGSSSNYMAHLDQMFSGVGVPTFDMGES- 320
QY 294 KMLEKCNQTPVRYVRVDMT-VPTAMWTGGQDWLGNPDEV-KMLSEVTLNLIYKHN--IP 349
Db 321 KNLKAYGQKLPQYNFTGIADVPYLFMSDDDLSTKQDEETLPAQLNSQVVOGSFRIE 380
QY 350 EWAHVDFIWGLDAPHMYNEIHLMOQE 378
Db 381 NYNHLHFINGTWASQVYVNTGIILODD 409

RESULT 11
T20480
hypotheical protein F01G10.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T20480
R:Hembry, C.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19280
A:Accession: T20480
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-426 <MIL>
A:Cross-references: EMBL:Z81055; PIDN:CAR02896.1; GSPDB:GN00022; CESP:F01G10.7
A:Experimental source: clone F01G10
C:Genetics:
A:Gene: CESP:F01G10.7
A:Map position: 4
A:Introns: 29/3; 64/1; 186/3; 347/2
C:Superfamily: triacylglycerol lipase, lingual

Query Match 31.4%; Score 651.5; DB 2; Length 426;
Best Local Similarity 36.8%; Pred. No. 5.7e-50;
Matches 136; Conservative 82; Mismatches 137; Indels 15; Gaps 8;

QY 7 AYDPEAFPMNISEIIHQGYPCPEYEVATEDGYLSVNRIP-RGLVQPKTSRPPVILQH 65
Db 20 AIDDECYMTVPPIGKHFGYSEVHLVTRTTDEYILELHRIPCKQNEKCDRSSKRPVFMQH 79
QY 66 GLVGGASNWSNLPNNSLGFILADAGFDVVMGNSRGNAMSRKHKTLISIDQDEFWAFSYDE 125
Db 80 GLLADGFSWIPNLANSAGFVPADAGFDIWIANSRGTPASQKHIGYGPENQKFWNFTWQ 139
QY 126 MARFDLPAVINFILOKTQGEKIYVYVSGQT-TMGFIAPSTWPELAQKIMYFALA 185
Db 140 MSEFDLTASVDVLKXETQGEFLYLGHSQGTIMFESRLAENPEFAKIRHFHALAPVATV 199
QY 186 KHAQSP---GTPKFLLLPDMIMKGLFGKKE---LQTRFLRQLVIYLCQVILDOIICSN 239
Db 200 SHIGGLFGLFGKFLTYAEI-----LLGRIPVSPISIPTRVQKIMYSMCSRPMQNICITD 255
QY 240 MLLGGFTNNMNSRASVVAHTLAGTSVQNLHWSQAVNSGELRAFDPWGSSETKXLEK 299
Db 256 IGFDIG-NEKMFNQSRGVYLCHTTPATSVKDLQHWIQLVKSQTSKFDYGTD-GNILEY 313
QY 300 NQPTVRYVRVDMT-VPTAMWTGGQDWLGNPDEV-KMLSEVTLNLIYK-NIPEWAVDFI 357
Db 314 GQTPPEPDLTQINTPTLYLWSRDDILADTDQIRDSILSKMKNKTAGSLELPHYSMDPV 373
QY 358 WGLDAPHRY 367
Db 374 FGTHAADFYL 383

RESULT 12

JT0949
 egg-specific protein - silkworm
 C;Species: Bombyx mori (silkworm)
 C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 26-Feb-1998
 C;Accession: JT0949, A28527
 R;Sato, Y.; Yamashita, O.
 submitted to JIPID, September 1991
 A;Reference number: JT0949
 A;Accession: JT0949
 A;Molecule type: DNA
 A;Residues: 1-559 <SAT>
 A;Experimental source: larva
 A;Note: this protein is a homotrimer
 R;Indrasith, L.S.; Sasaki, T.; Yamashita, O.
 J. Biol. Chem. 263, 1045-1051, 1988
 A;Title: A unique protease responsible for selective degradation of a yolk protein in B.
 A;Reference number: A28527; PMID:88087166; PMID:3275655
 A;Accession: A28527
 A;Molecule type: protein
 A;Residues: 19-40;133-144, 'T', 146-152;229-248 <IND>
 C;Genetics:
 A;Map position: 19
 A;Keywords: egg yolk; homotrimer
 F;132-133/Cleavage site: Lys-Asn (egg 30.5K cysteine proteinase) #status experimental
 F;228-229/Cleavage site: Arg-Asp (egg 30.5K cysteine proteinase) #status experimental

Query Match 27.6%; Score 574; DB 2; Length 559;
 Best Local Similarity 33.5%; Pred. No. 6.6e-43;
 Matches 130; Conservative 81; Mismatches 149; Indels 28; Gaps 8;

QY 5 TKAVDPEAF-MNISEIIHQGYCEYEVEDGYILSNRIPRGLVQPKTKGSRPVLL 63
 DB 180 TQRENNEDEFLNATELLKQYVVEHTVATDDGYHLTVLRIP-PHTRDDKCKPVAL 238
 QY 64 QHGLVGASNIWNLNPNNSGFLTADAGPDVVMGNSRGNAWRKHKHTLSIDODEFWAFSY 123
 DB 239 MHGLGASADDWLLMGPSKSLAYLWLSADGYDVLGNVGRNGKYSRSHVKEPALNDFWKF 298
 QY 124 DENARFDLPVINFILQKQOEKIYVYVGSQGTMGFIAPFTMPPELAQIKMYFALAPIA 183
 DB 299 DEIALHDLPAIDHVLDISQERLHYIGHSQGAATTFALMSEQPSYNEKIVSMHALSPIV 358
 QY 184 TVKHAKS-----PGTKFL-LLPDMMIKGLFKCKEFLYQTRFLRLQVLYLQGVILDO- 234
 DB 359 YMYVRSPLFRMTAPSKFYQYHQVGHGAFEPKHLIET-----FGAACREKL 409
 QY 235 -----ICSNIMLLGGFNTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNSGELRAFDWG 290
 DB 410 GCRHVCNMLNVISGINVYNQDADIVPVVMAHLPAQTSARVMKQYQGVASHDFRKYNG 469
 QY 291 SETKYLEKQKQPTVRVVRVDMVTPTAMWTGGQDWLSPNPDVKNLLSEVNTLIYHKNIEP 350
 DB 470 AET-NMKVYGASPEPSSDLKSVAPNLYSHDAWLAHPKQVKELOENTLPNVKQSFVEPE 528
 QY 351 WAH-----VDFTWGLDAPHRMNEIHLMO 375
 DB 529 QQHTDLDLDFQSKKAPDVTYOKLMENMO 556

RESULT 13

T39540
 triglyceride lipase-cholesterol esterase - fission yeast (Schizosaccharomyces pombe)
 C;Species: Schizosaccharomyces pombe
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
 C;Accession: T39540
 R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Beck, A.; Reinhardt, R.
 submitted to the EMBL Data Library, February 1998
 A;Reference number: Z21862
 A;Accession: T39540
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA

A;Residues: 1-443 <WOO>
 A;Cross-references: EMBL:AL021748; PIDN:CAA16863.1; GSPDB:GN00067; SPDB:SPBCL6A3.12c
 A;Experimental source: strain 972h-; cosmids c16A3
 C;Genetics:
 A;Gene: SPDB:SPBCL6A3.12c
 A;Map position: 2
 A;Introns: 11/2; 152/3; 282/1; 398/3
 C;Superfamily: triacylglycerol lipase, lingual

Query Match 23.7%; Score 491; DB 2; Length 443;
 Best Local Similarity 30.7%; Pred. No. 1.2e-35;
 Matches 117; Conservative 81; Mismatches 135; Indels 48; Gaps 11;

QY 15 NISEIIHQGYCEYEVEDGYILSNRIPRGLVQPKTKGSRPVLLQHLGLVGASN 73
 DB 74 NIVEICEAFGYRVEEHLVTRQDNFILLCHRI-----THPKSQSKREVYVCHGLMTNSEL 129
 QY 74 WIS-NLPNNSGFLTADAGPDVVMGNSRGNAWRKHKHTLSIDODEFWAFSYDEMAFIDL 132
 DB 130 WVAVESERSLPLFVLESYDGVVLMGNRGNKYSRKHITYPKDEEFNFSLDMAFIDIP 189
 QY 133 AVINFILQKQOEKIYVYVGSQGTMGFIAPFTMPPELAQIKMYFALAPIATVYKHA 192
 DB 190 DTVDYILRETGREKLYIGFSQGTQAQMAALSINPLNDKVNIFIGLAPAYAPKGF 249
 QY 193 TKFLLL--PDMMIKGLFKCKEFLYQTRFLRLQVLYLQGVILDOICSNIMLLGGFNT 250
 DB 250 VDYIVKVPKIMYH-LFGRKCLLPSTFWQ-----ICYPPIFVKIVDVSLKILFN 304
 QY 251 MNMSRASVYAAHTLAGTSVQNLHWSQAVNSGELRAFD-----WGSE-----TK 294
 DB 305 ISLNQKLCGVAHLYSFSSVSVVHMLQIKNCTFQLYDDMALLAGVGRHYQVPLPPTN 364
 QY 295 NLEKQKQPTVRVVRVDMVTPTAMWTGGQDWLSPNPDVKNLLSEVNTLIYHKNIE 354
 DB 365 NI-KC-----FMLILMGKQDTLINMEVMTALPFAKEV---SIAHYEHL 405
 QY 355 DFIWGLDAPHRMNEIHLMO 375
 DB 406 DFLMGQDVKEVEFPVVIDALK 426

RESULT 14

T31611
 hypothetical protein Y508A.g - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C;Accession: T31611
 R;Steward, C.
 submitted to the EMBL Data Library, September 1999
 A;Reference number: Z21047
 A;Accession: T31611
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-1585 <WIL>
 A;Cross-references: EMBL:AL117200; NID:el549770; PIDN:CAB55050.1; CESP:Y508A.9
 A;Experimental source: clone Y508A
 C;Genetics:
 A;Gene: CESP:Y508A.9
 A;Introns: 25/3; 60/1; 133/2; 217/3; 270/3; 337/2; 400/1; 746/2

Query Match 23.0%; Score 478.5; DB 2; Length 1585;
 Best Local Similarity 30.6%; Pred. No. 9e-34;
 Matches 118; Conservative 67; Mismatches 172; Indels 29; Gaps 9;

QY 11 EAFNISEIIHQGYCEYEVEDGYILSNRIPRGLVQPKTKG--SRPVLLQHLGLV 68
 DB 20 EETLDAADTISHYGYVEKHVTTDDGYTVQLQRIPVGRDDRSILGCKEPVVFVHGLF 79
 QY 69 CGASNISNLPNNSGFLTADAGPDVVMGNSRGNAWRKHKHTLSIDODEFWAFSYDEMA 128
 DB 80 GSSYHFLNLPSSQAAYIFADAGVDVWLGIRNGTEYGLNHTSPSTNGVNFNFSLEYHSH 139

Search completed: February 19, 2004, 14:23:08
Job time : 15.4317 secs

```
QY 129 FDLPAVINFIQKQGEKIYVYGSQGTMTMGFIATSTMPEL-----ACKIKWYFALAPIAT 184
Db 140 YDLRQQLYYVLDYTRHESLTVVGHSSQTA---VMFARLAADVWQSKIEVFFALGFTAG 196
QY 185 VKHAKSPGTFFLLPDMIKGLPKGKFLYQTRP-----LRQLVYILCGQVILDO 234
Db 197 FMKPLMP---FTLLEENYLOALI---QFALDGRFGILPVEIPRAIASKFADFCSSKFPTF 250
QY 235 ICNIMILLGGFNT--NNMNSRASVYAAHTLAGTSVONILHWSQAVNSGELRAFDWGSSET 293
Db 251 LCSAGPKVAAGIETLQGVNDSRPIILSHSPSATSTNLHLLHWQIFKYHLERLDLGT-A 309
QY 294 KNLEKCMQPTPVVRVDMVTPTAMWTGGQDMLSNPEDVQMLSEV---TNLIYHKNIPEW 351
Db 310 RNLIAQCKDAPRLIEIGNIIAQTLIFYSKDDQITDEVDVREIIMKQMGPLIESYLDLHF 369
QY 352 AHVDPIWGLDAPHRMVNEIHLMOOE 377
Db 370 THDFILGLRATDEVYKPIVYRIYKE 395

RESULT 15
T41053
triglyceride lipase-cholesterol esterase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2003
C:Accession: T41053
R:Halbert, H.; Duesterhoeft, A.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, August 1998
A:Reference number: Z21967
A:Accession: T41053
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-467 <HIL>
A:Cross-references: EMBL:AL031324; PIDN:CAA20447.1; GSPDB:GN00068; SPDB:SPCC1672.09
A:Experimental source: strain 972h-; cosmid c1672
C:Genetics:
A:Gene: SPDB:SPCC1672.09
A:Map position: 3
A:Introns: 12/2; 164/3; 294/1
C:Superfamily: triacylglycerol lipase, lingual

Query Match 22.9%; Score 475; DB 2; Length 467;
Best Local Similarity 31.2%; Pred. No. 3.4e-34;
Matches 125; Conservative 68; Mismatches 158; Indels 50; Gaps 12;

QY 6 KAVDPEAFMNISEIIHQGYPCEEYEVATEDGYILSVNRIPR-----GLVQPKYKTSRP 59
Db 73 KAIHDA--DIRECKISGYTVEDHLVRTEDDYILCHRSKOSPGRIGSPFKCL--P 127
QY 60 VVLLQHGLVGGASNI--PNNLSGLFILDAGFDVWNGSRGNWRSRKHKTLSIDQDEF 118
Db 128 VVYCHGGLMNSVWVCNVDPNCLVFDLVNKGVDVWLGNNRGKYSRQHLRFDSTDKF 187
QY 119 WAFSYDEMARFDLPAVINFIQKQGEKIYVYGSQGTMTMGFIATSTMPELAQIKMYFA 178
Db 188 WDFSIDDPFAQYDIPDTIDYILKTSQTKLTIVYGSQGTAFAPASLSIHPILNDKINSLIA 247
QY 179 LAPATVNHAKSPGKTKFLLPDMWIKG-----LFGKGFYOTRFLRQLVYILCGQVI 231
Db 248 LAP-----AISPGLHNRVVDAPVKARPSILFFLFGKSLPSAGFWQSFL-----APKF 297
QY 232 LDQICSNIMLLGGFNTNNMNSRASVYAAHTLAGTSVONILHWSQAVNSGELRAFD--- 288
Db 298 FDAVLAYCLSLQFNWSCQNISSYQRLVSFAHLSYTSYKCLVHWFQIMRSAEFMYDNDQ 357
QY 289 WGSE--TKNLEKCMQPTPVVRVDMVTPTAMWTGGQDMLSNPEDVQMLSEVTLNLIYHK 346
Db 358 LGHYDFLYKYKAAKEPT-----NNIRTPYLIWGGSDSL---VDIQAMNLPAEVEHV 408
QY 347 NIPWAHVDFIWG-----LDAPHRMYNEIHLMOOEETN 380
Db 409 KVDSYEHLDMIWADTVKDYVPPVPLRLRLDIHPPHEEND 449
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2004, 14:17:14 ; Search time 27.8229 Seconds
(without alignments)
3617.186 Million cell updates/sec

Title: US-10-042-431-49

Perfect score: 2076

Sequence: 1 VHMPTKAVDEAFNWESEII.....IHLMQEETNLSQRCSEAVL 390

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 23:*

- 1: sp_archea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1991	95.9	374	Q96LG2	Q96LG2 homo sapien
2	1848.5	89.0	422	Q8K2A6	Q8K2A6 mus musculus
3	1844.5	88.8	422	Q8BJ28	Q8BJ28 mus musculus
4	1219.5	58.7	397	Q8C3G7	Q8C3G7 mus musculus
5	1130	54.4	395	Q8D798	Q8D798 mus musculus
6	1130	54.4	395	Q8D6Q6	Q8D6Q6 mus musculus
7	1130	54.4	395	Q8D6P3	Q8D6P3 mus musculus
8	1130	54.4	395	Q8D6X0	Q8D6X0 mus musculus
9	1129	54.4	395	Q8C3P7	Q8C3P7 mus musculus
10	1129	54.4	395	Q8C3P8	Q8C3P8 mus musculus
11	1127	54.3	395	Q8D6L9	Q8D6L9 mus musculus
12	1125	54.2	395	Q8D6T5	Q8D6T5 mus musculus
13	1124	54.1	395	Q8D767	Q8D767 mus musculus
14	1123	54.1	395	Q8D7C5	Q8D7C5 mus musculus
15	1122	54.0	395	Q8D6L1	Q8D6L1 mus musculus
16	1122	54.0	395	Q8D766	Q8D766 mus musculus

17	1121	54.0	395	11	Q9D6Q3	Q9D6Q3 mus musculus
18	1120	53.9	395	11	Q9D760	Q9D760 mus musculus
19	1120	53.9	395	11	Q9D6S5	Q9D6S5 mus musculus
20	1118	53.9	395	11	Q9D796	Q9D796 mus musculus
21	1115	53.7	395	11	Q9D6N8	Q9D6N8 mus musculus
22	1112	53.6	395	11	Q9D6Q2	Q9D6Q2 mus musculus
23	1111	53.5	395	11	Q9D6M9	Q9D6M9 mus musculus
24	1095	52.7	398	11	Q8BM14	Q8BM14 mus musculus
25	829.5	40.0	275	13	Q8AYR8	Q8AYR8 pseudopleur
26	755.5	36.4	684	5	Q95XV1	Q95XV1 caenorhabdi
27	747	36.0	411	5	Q20449	Q20449 caenorhabdi
28	744.5	35.9	403	5	Q61866	Q61866 caenorhabdi
29	727	35.0	405	5	Q93789	Q93789 caenorhabdi
30	707.5	34.1	404	5	Q16956	Q16956 caenorhabdi
31	691.5	33.3	411	5	Q94252	Q94252 caenorhabdi
32	660.5	31.8	351	5	Q95X33	Q95X33 caenorhabdi
33	655.5	31.6	434	5	Q9VKT9	Q9VKT9 drosophila
34	654	31.5	422	5	Q9VKT8	Q9VKT8 drosophila
35	651.5	31.4	426	5	Q17768	Q17768 caenorhabdi
36	650.5	31.3	531	5	Q9VKS5	Q9VKS5 drosophila
37	636	30.6	398	5	Q9V796	Q9V796 drosophila
38	625.5	30.1	616	5	Q77107	Q77107 plodia inte
39	602	29.0	429	5	Q8MMX1	Q8MMX1 dictyosteli
40	580.5	28.0	421	5	Q9VBK6	Q9VBK6 drosophila
41	574	27.6	559	5	Q17219	Q17219 bombyx mori
42	567.5	27.3	457	5	Q9VKT2	Q9VKT2 drosophila
43	567.5	27.3	457	5	Q8T3X7	Q8T3X7 drosophila
44	559	26.9	399	5	Q9VPE9	Q9VPE9 drosophila
45	535	25.8	416	5	Q9VQG5	Q9VQG5 drosophila

ALIGNMENTS

RESULT 1

Q96LG2 ID Q96LG2 PRELIMINARY; PRT; 374 AA.
AC AC Q96LG2;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE BA30415.1 (Novel lipase) (Fragment).
GN BA30415.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bray-Allen S.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AL353113; CAC78754.1; -
DR InterPro; IPR000073; A/b hydrolase.
DR InterPro; IPR006693; abhydro_lipase.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser_estr_site.
DR Pfam; PF00561; abhydro_lipase; 1.
DR Pfam; PF04083; abhydro_lipase; 1.
DR PROSITE; PS00120; LIPASE_SER; 1.
FT NON TER 1
SQ SEQUENCE 374 AA; 42371 MW; 5E7220A889437337 CRC64;

Query Match 95.9%; Score 1991; DB 4; Length 374;

Best Local Similarity 100.0%; Pred. No. 1.3e-175;

Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 SEIIHQGPCEYEYVATEDGYLSVNRIPRLGVOPKKTGSRPVLLQHLGVGGASNMIS 76

Db 1 SEIIHQGPCEYEYVATEDGYLSVNRIPRLGVOPKKTGSRPVLLQHLGVGGASNMIS 60

QY 77 NLPNNSLGFILADAGFDVMMGNMGRNNAWSRKHTLSIQDDEFWAFSYDEMAFDPDLAVIN 136

Db 61 NLPNNSLGFILADAGFDVMMGNMGRNNAWSRKHTLSIQDDEFWAFSYDEMAFDPDLAVIN 120

QY 137 FILQKTGQEKIYVGVSGTGTGFIAPSTWPELAOKIMYFALAPIATVTHAKSPGTKFL 196
 Db 121 FILQKTGQEKIYVGVSGTGTGFIAPSTWPELAOKIMYFALAPIATVTHAKSPGTKFL 180
 QY 197 LLPDMIKGLFGKKEFLYQTRFLRQLVYLCOGVILDOICSNIMLLGGFNTNNMGR 256
 Db 181 LLPDMIKGLFGKKEFLYQTRFLRQLVYLCOGVILDOICSNIMLLGGFNTNNMGR 240
 QY 257 SVTAHTLAGTSVONILHWSQVNSGELRAFDSGSETKLEKNCQPTPVRYVRDVTPT 316
 Db 241 SVTAHTLAGTSVONILHWSQVNSGELRAFDSGSETKLEKNCQPTPVRYVRDVTPT 300
 QY 317 AMWTGGQDMLSNPDEKVMKLLSEVNTLIYHKNIPWAHVDVFWGLDAPHRYNEIHLMOQ 376
 Db 301 AMWTGGQDMLSNPDEKVMKLLSEVNTLIYHKNIPWAHVDVFWGLDAPHRYNEIHLMOQ 360
 QY 377 EETNLSQGRCEAVL 390
 Db 361 EETNLSQGRCEAVL 374

RESULT 2

Q8K2A6 PRELIMINARY; PRT; 422 AA.
 AC Q8K2A6
 DT 01-OCT-2002 (TREMELrel. 22, Created)
 DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
 DE RIKEN CDNA 4632427C23 gene (BA30415.1).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Mammary gland, and Vagina;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,

RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 DR EMBL; BC031933; AAH31933.1; -
 DR EMBL; AK037031; BAC29899.1; -
 DR EMBL; AK085719; BAC39517.1; -
 DR InterPro; IPR000073; A/b hydrolase.
 DR InterPro; IPR006693; abhydro_lipase.
 DR InterPro; IPR000734; Lipase.
 DR InterPro; IPR000379; Ser esterase.
 DR Pfam; PF00561; abhydro_lipase; 1.
 DR Pfam; PF04083; abhydro_lipase; 1.
 DR PROSITE; PS00120; LIPASE_SER; 1.
 SQ SEQUENCE 422 AA; 48253 MW; C6EFEC140963B051 CRC64;

Query Match 89.0%; Score 1848.5; DB 11; Length 422;
 Best Local Similarity 88.4%; Pred. No. 2.4e-162;
 Matches 344; Conservative 20; Mismatches 24; Indels 1; Gaps 1;

QY 2 HMPKAVDPPEAFMNIISIIHQHGYPCBEYEVATEDGVILSVNRIIPRGVOPKKTGSRPVV 61
 Db 35 HLPKAAADPEAFMNVSEIIKHGYPSEYEVATEDGVILSVNRIIPRGVOPKKTGSRPVV 94
 QY 62 LLOHGLVGASNWI SNLPNNSLGFILADAGFDVVMGNSRGNSRKHKTLSIDQDEFWAF 121
 Db 95 LLOHGLVGASNWI SNLPNNSLGFILADAGFDVVMGNSRGNSRKHKTLSIDQDEFWAF 154
 QY 122 SYDEMARFDPVAVINFILOKTGQEKIYVGVSGTGTGFIAPSTWPELAOKIMYFALAP 181

Db 155 SYDEMARFDPVAVINFILOKTGQEKIYVGVSGTGTGFIAPSTWPELAOKIMYFALAP 214
 QY 182 IATVTHAKSPGTKFLLPDMIKGLFGKKEFLYQTRFLRQLVYLCOGVILDOICSNIML 241
 Db 215 IATVTHAKSPGTKFLLPDMIKGLFGKKEFLYQTRFLRQLVYLCOGVILDOICSNIML 274
 QY 242 LLOGFNTNNMGRASVYAHTLAGTSVONILHWSQVNSGELRAFDSGSETKLEKNCQ 301
 Db 275 LLOGFNTNNMGRASVYAHTLAGTSVONILHWSQVNSGELRAFDSGSETKLEKNCQ 334
 QY 302 PTPVRYVRDVTPTAMWTGGQDMLSNPDEKVMKLLSEVNTLIYHKNIPWAHVDVFWGLD 361
 Db 335 PTPVRYVRDVTPTAMWTGGQDMLSNPDEKVMKLLSEVNTLIYHKNIPWAHVDVFWGLD 394
 QY 362 APHRMYNEIHLMOQEEETNLSQGRCEAVL 390
 Db 395 APQRVYNEIHLMAKQ-EPNLPQGTCTRVKL 422

RESULT 3

Q8BU28 PRELIMINARY; PRT; 422 AA.
 AC Q8BU28;
 DT 01-MAR-2003 (TREMELrel. 23, Created)
 DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
 DE BA30415.1
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Skin;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,

RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 DR EMBL; AK037214; BAC29757.1; -
 SQ SEQUENCE 422 AA; 48267 MW; 5F2522D0E27050C8 CRC64;

Query Match 88.8%; Score 1844.5; DB 11; Length 422;
 Best Local Similarity 88.2%; Pred. No. 5.7e-162;
 Matches 343; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

QY 2 HMPKAVDPPEAFMNIISIIHQHGYPCBEYEVATEDGVILSVNRIIPRGVOPKKTGSRPVV 61
 Db 35 HLPKAAADPEAFMNVSEIIKHGYPSEYEVATEDGVILSVNRIIPRGVOPKKTGSRPVV 94
 QY 62 LLOHGLVGASNWI SNLPNNSLGFILADAGFDVVMGNSRGNSRKHKTLSIDQDEFWAF 121
 Db 95 LLOHGLVGASNWI SNLPNNSLGFILADAGFDVVMGNSRGNSRKHKTLSIDQDEFWAF 154
 QY 122 SYDEMARFDPVAVINFILOKTGQEKIYVGVSGTGTGFIAPSTWPELAOKIMYFALAP 181
 Db 155 SYDEMARFDPVAVINFILOKTGQEKIYVGVSGTGTGFIAPSTWPELAOKIMYFALAP 214
 QY 182 IATVTHAKSPGTKFLLPDMIKGLFGKKEFLYQTRFLRQLVYLCOGVILDOICSNIML 241
 Db 215 IATVTHAKSPGTKFLLPDMIKGLFGKKEFLYQTRFLRQLVYLCOGVILDOICSNIML 274
 QY 242 LLOGFNTNNMGRASVYAHTLAGTSVONILHWSQVNSGELRAFDSGSETKLEKNCQ 301
 Db 275 LLOGFNTNNMGRASVYAHTLAGTSVONILHWSQVNSGELRAFDSGSETKLEKNCQ 334
 QY 302 PTPVRYVRDVTPTAMWTGGQDMLSNPDEKVMKLLSEVNTLIYHKNIPWAHVDVFWGLD 361
 Db 335 PTPVRYVRDVTPTAMWTGGQDMLSNPDEKVMKLLSEVNTLIYHKNIPWAHVDVFWGLD 394
 QY 362 APHRMYNEIHLMOQEEETNLSQGRCEAVL 390

```
Db 395 APOQVYNEIHLKQ-EPNLPQGTCTVKL 422
RESULT 4
Q8C2G7 PRELIMINARY; PRT; 397 AA.
AC Q8C2G7;
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Lysosomal acid lipase 1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK888659; BAC40484.1; -.
SQ SEQUENCE 397 AA; 45325 MW; 866B123A1A058134 CRC64;

Query Match 58.7%; Score 1219.5; DB 11; Length 397;
Best Local Similarity 58.4%; Pred. No. 3.5e-104;
Matches 220; Conservative 66; Mismatches 88; Indels 3; Gaps 1;

QY 3 MPT---KAVDPKPNVISEIIHQGYPCPEYEVATEDGYILSVNRIPIRGVLQPKTGRSP 59
Db 19 VPTGTVAVDPEVNNVTEIIMRWGYPGEHSVLTGDTYILSIHPIRGKRNHFGKPRP 78
QY 60 VVYLQHLGVGASNIWNLNPNLSGLFILADAGFDVWNGSRGNASRKHHTLSIDQDEFW 119
Db 79 VVYLQHLGLADSSNWTNIDNSSLGFLILADAGFDVWNGSRGNTWSLKHHTLSVSDQEFW 138
QY 120 AFSVDEMARFPLPAVINFILOKTOEKIYVYVYSGTMTGFIAPSTMPDLAKIKMYFAL 179
Db 139 AFSVDEMAKYLPAVINIYLNKTOEQIYVYVYVYSGTGTIGFIAPSQMPELAKIKMYFLV 198
QY 180 APIATVHKASFGTKFLLPDMITKGLFGKKEFLYQTRFLRQLVYILCGOVLIDQICNI 239
Db 199 APVLSLNFASGLQLGLPDLKDMFGKQFLPQSAMLKWLSHVCTHVIMKELCANV 258
QY 240 MLLGGFNWNNMSRASVYAAHTLAGTSVQNIILHWSQAVNSGELRAPDWGSETKNLEK 299
Db 259 FFLCGFNEKLNLSRVDVYTHCPAGTSVQNMHLHWGQVFKYKLOAPDWGSEKKNFYHY 318
QY 300 NQPTPVRYRDMTPTAMWTGGODWLSNPEDVKMLLSEVNTLIYHKNIPEWAHVDFIWG 359
Db 319 NQSPFPNINKNMLPTALNSGGDLADINDIITLTQPKLYHKNIPEWHDLDPIWG 378
QY 360 LDAPHMYNEIHLMOQ 376
Db 379 LDAPKLYDEIISLMKX 395

RESULT 5
Q9D798 PRELIMINARY; PRT; 395 AA.
AC Q9D798;
DT 01-JUN-2001 (TREMELrel. 17, Created)
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE 2310051B2IRik protein.
GN 2310051B2IRIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Mateu Y., Nikaudo I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaudo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
DR EMBL; AK009431; BAB26283.1; -.
DR HSSP; P07098; IHLG.
DR MGD; MGI:1914967; 2310051B2IRik.
DR InterPro; IPR000073; A/b hydrolase.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser esterase site.
DR Pfam; PF00561; abhydrolase_1_site.
DR PROSITE; PS00120; LIPASE SER; 1.
SQ SEQUENCE 395 AA; 44603 MW; D3FDB8B6FA671E3E CRC64;

Query Match 54.4%; Score 1130; DB 11; Length 395;
Best Local Similarity 56.2%; Pred. No. 6.7e-96;
Matches 209; Conservative 64; Mismatches 93; Indels 6; Gaps 2;

QY 9 DPEAFNMTSEIIHQGYPCPEYEVATEDGYILSVNRIPIRGVLQPKTGRSPVLLQHLGV 68
Db 27 NPEANNVSQMITIYWGYPSEYEVVTEGYILGVTRIFYPKKNSENIGKRPVAYLQHLGI 86
QY 69 GGASNIWNLNPNLSGLFILADAGFDVWNGSRGNASRKHHTLSIDQDEFWAFSDENAR 128
Db 87 ASATNWTNLPNLSLAFILADAGYDVLGNSRGNTWSKNNVYSPDSVEFWAFSDENAK 146
QY 129 FDLPAVINFILOKTOEKIYVYVYVYSGTMTGFIAPSTMPDLAKIKMYFALAPATYKHA 188
Db 147 YDLPATIDFIVOKTOEKIHYVYVYVYSGTGTIGFIAPSTMPALAKIKRFPALAPATYK 206
QY 189 KSPGTFKLLPDMITKGLFGKKEFLYQTRFLRQLVYILCGOVLIDQICSNIMLLGG 245
Db 207 ESPFKKISLIPKFLKLVFGNKFMPHNYLDQFLGTEV---CSRELLDLLCSNALFIFCG 263
QY 246 FNTNNMNSRASVYAAHTLAGTSVQNIILHWSQAVNSGELRAPDWGSETKNLEKQPTPV 305
Db 264 FDKNLNLSRDLVYLGHPAGTSTQDLFWAQLAKSGKLOAYNNGSPLOQNMHLHYNQKTPP 323
QY 306 RYKVRDMTPTAMWTGGODWLSNPEDVKMLLSEVNTLIYHKNIPEWAHVDFIWLGDAPHR 365
Db 324 YDVSNTVPIAVNWGGHDLADPDQVAMLLPKFLNLYHKELLYHNLDFIWMADAPQE 383
QY 366 MYNEIHLMOQ 377
Db 384 VYNEIVTMAED 395

RESULT 6
Q9D6Q6 PRELIMINARY; PRT; 395 AA.
AC Q9D6Q6;
DT 01-JUN-2001 (TREMELrel. 17, Created)
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
```

[illegible]


```

DR EMBL; AK010125; BAB2671.1; -.
DR HSP; P07098; 1HLG
DR MGD; MGI:1914967; 2310051B21Rik.
DR InterPro; IPR000073; A/B_hydrolase.
DR InterPro; IPR000379; Ser_estra_site.
DR Pfam; PF00561; abhydrolase; 1.
DR PROSITE; PS00120; LIPASE SER; 1.
SQ SEQUENCE 395 AA; 44637 MW; D3F96B65EA671B34 CRC64;

Query Match 54.4%; Score 1129; DB 11; Length 395;
Best Local Similarity 56.2%; Pred. No. 8.3e-96;
Matches 209; Conservative 64; Mismatches 93; Indels 6; Gaps 2;

QY 9 DPEAFMISIIHQGYPCBEYEVATEDGYILSVNRIPLRGVOPKKTGSRPVVLLQHLV 68
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
27 NPEANMVSQMITWGYPSSEYEVATEDGYILGVYRIPYKKNSENIKRPVAYLQHLI 86
QY 69 GGASNWTISLNNLSGLIFILADAGFDVWNGSRGNVSRKHTLSIDQDFWAFSYDEMAR 128
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
87 ASATNWTISLNNLSGLIFILADAGFDVWNGSRGNVSRKHTLSIDQDFWAFSYDEMAR 146
QY 129 FDLPAVINFLIQTQGEKIYVGVSGQTTMGFTAFSTMPBELAKIKMYFALAPIATVCHA 188
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
147 YDLPATIDFIVQKTQGEKIYVGVSGQTTMGFTAFSTMPBELAKIKMYFALAPIATVCHA 206
QY 189 KSPQTKFLLPDDMKIGLFGKKEPL---YOTRFLRQLVYVLCQVILDOICSNIMLLGG 245
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
207 ESPFKKISLIPKFLKVIIFGNKMFPHNYLDQFLGTEV---CSRELLDLCSNALFIFCG 263
QY 246 FNTNNMNSRASVVAHTLAGTSVQNTLHWSQVNSGELRAFDPWGSSTKLEKCNQPTPV 305
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
264 FDKQLNVSRFDVYLGHNPAQTSTQDLFWAQLAKSGKQLQYNGSPQLQMLHYNQKTPP 323
QY 306 RYVRDMVTPTAMTGGQDMLSNPDEVKMLLSEVTNLIYHNKIPENAHVDFIWLGDAPHR 365
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
324 YDVSAMTPTAVWNGGHIDLPQDVAMLLPKLPNLLYHKEILPYNHLDFIWAMDAPOE 383
QY 366 MYNEIHLMOQE 377
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
384 VYNEIVTMMAED 395

RESULT 10
Q9CPE8 ID Q9CPE8 PRELIMINARY; PRT; 395 AA.
AC Q9CPE8
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE 2310051B21Rik protein.
GN 2310051B21Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSU=Tongue;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Iizawa M., Nishi K., Kiyosawa H., Konno H., Adachi J., Fukuda S.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stauber F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boifelli D., Boylston N., Carrinchi P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

```

```

[1]
RN  SEQUENCE FROM N.A.
RP  STRAIN=C57BL/6J; TISSUE=Tongue;
RC  MEDLINE=21088660; PubMed=1217851;
RX  Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA  Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA  Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA  Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA  Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA  Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA  Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA  Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA  Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsch G.,
RA  Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.P.,
RA  Brownstein M.J., But C., Fletcher C., Fujita M., Gariboldi M.,
RA  Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA  Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA  Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA  Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA  Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA  Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsuki S.,
RA  Hayashizaki Y.;
RT  "Functional annotation of a full-length mouse cDNA collection.";
RL  Nature 409:685-690(2001).
DR  EMBL; AK009990; BAB26629.1; -.
DR  HSP; P07098; 1HLG.
DR  MGD; MGI:1914967; 2310051B21Rik.
DR  InterPro; IPR000073; A/b_hydrolase.
DR  InterPro; IPR000734; Lipase.
DR  InterPro; IPR000379; Ser_estrse_site.
DR  Pfam; PF00561; abhydrolase_1.
DR  PROSITE; P800120; LIPASE_SER; 1.
SQ  SEQUENCE 395 AA; 44623 MW; D3F96DB83161C3EF CRC64;

Query Match      54.2%; Score 1125; DB 11; Length 395;
Best Local Similarity 55.9%; Pred. No. 1.9e-95;
Matches 208; Conservative 65; Mismatches 93; Indels 6; Gaps 2

QY      9  DPEAFNLTSEIIHQGYPCPEYEVATEDGYILSVNRIPRGLVQPKKTGRPVVLLQHGVLV 88
Db      27  NPEANMYSQMTIWTGYPSEYEVVEDGYILGYTRIPYGGKNSENIGKRPVAYLQHGILI 86

QY      69  GGASNIWISLNPNSLGFTLADAGPDVWGMNSRGNANRSKHTTUSIDODZFWAFSYDEMAR 128
Db      87  ASATNTWITLNPNSLAFITADAGDYVLGNISRGNTWRSKNVYSPDSVFWAFSFDDEMAK 146

QY      129  FDLPAVINFILOKTQGEKIYVGVYSQGTWGTGFTAFSTMPBELAQIKWYFALAPIATVKA 188
Db      147  YDLPATIDFIYOKTQGEKIHTVGHISQGTITIGTAFSTNPALAKIKRFTALPAPVATVKT 206

QY      189  KSPGTFKLLLPDMIMKGLFGKKEFL---YQTRFLRQLVLYLVCQGVILDOICSNIMLLGG 245
Db      207  ESFPFKKLSILPKFKLLKVFIGNKMPHNYLDQFLGTEV---CSRELLDLLCSNALPFIQCG 263

QY      246  FWTNNMWSRASVYAAHTLAGTSQNTILWSQAVNSGELPAFQWGSFTKNLEKCKQPTPV 305
Db      264  FDKKLNLSRFDVYLGHNPAGTSDQDLFWQAQAKSGKQAYNWGSPQMLHLHYNQKSP 323

QY      306  RYRVDMTVPTAMTGGQDMLNSPDEDVKMLLSEVTNLIYHKNTPENAHYDFIWGLDAPHR 365
Db      324  YDVSAMTVPLAVWNGGHIDLADPDQVAMLLPKLPLNLLYHKEILPYNHLDFIWMDAPQE 383

QY      366  MTNEIHLMQOE 377
Db      384  VTNEIVTWMMAED 395

RESULT 13
Q9D767
ID   Q9D767          PRELIMINARY;      PRT;       395 AA.
AC   Q9D767;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

```

RESULT 14					
ID	Q9D7C5	PRELIMINARY;	PRT;	395 AA.	
AC	Q9D7C5;				
DT	01-JUN-2001 (TrEMBLrel. 17, Created)				
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)				
DE	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)				
DE	2310051B21RIK protein.				
GN	2310051B21RIK				
OS	Mus musculus (Mus.)				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
NB	NCBI_TaxId=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=C57BL/6J; TISSUE=Tongue;				
RX	MEDLINE=21085660; PubMed=11217851;				
RA	Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,				
RA	Arakawa T., Hara A., Fukunishi Y., Komio H., Adachi J., Fukuda S.,				
RA	Aizawa K., Izawa M., Nishi K., Kiyoosawa H., Kondo S., Yamada I.,				
RA	Saito T., Okazaki Y., Gotojori T., Bono H., Kasukawa T., Saito R.,				
RA	Kadota K., Matsuda H.A., Asburner M., Batalov S., Casavant T.,				
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,				
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,				
RA	Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,				
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,				
RA	Blake J., Buffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,				
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Kamiya M., Lee N.H.,				
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,				
RA	Livons P., Marchionni L., Mashima J., Mazzarelli J., Momabaets P.,				
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,				
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,				
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,				
RA	Wyshaw-Boris A., Yoshida K., Haegawa Y., Kawaji H., Kohtsuki S.,				
RA	Hayashizaki Y.;				
RT	"Functional annotation of a full-length mouse cDNA collection.";				
RL	Nature 409:685-690(2001).				
DR	EMBL; AK009359; BAB26240.1; -				
DR	HSP; P07098; IHLG.				
DR	MGI; MG11914967; 2310051B21RIK.				
DR	InterPro; IPRO00073; A/b hydrolase.				
DR	InterPro; IPRO00734; Lipase.				
DR	InterPro; IPRO00379; Ser esterase site.				
DR	Fram; PF00561; abhydrolase; i.				
DR	PROSITE; PS00120; LIPASE_SER; 1.				
SQ	SEQUENCE 395 AA; 44748 MW; 3ED3A992458DE058 CRC64;				
Query Match 54.1%; Score 1123; DB 11; Length 395;					
Best Local Similarity 55.9%; Pred. No. 3e-95;					
Matches 208; Conservative 63; Mismatches 95; Indels 6; Gaps 2;					
Qy	9	DPEAFNMISIIHQGYPCEYEYVATGGVLTLSVNIRPGVLGPVKTKGRPVVLQHLGV	68		
Db	27	NPEANMVMSOMITWYGPPSYSEYEVTDGGVILGYRIPYKKNSNIKRPVAYLQHGLI	86		
Qy	69	GGASNNVISLNPNNSLGFILADAGPDVVMGNSRGNAWSRKETLSIDQDFWAFSDEVAR	128		
Db	87	ASAKNWITNLPNNSLAIFILADAGDYVWLNGSRGTWRKNVYSPDSVEFWAFSDEMAK	146		
Qy	129	FDLPVAVINFLQKTQSKIYYVGSCQTGWGTFAPSTMPELAQIKMYFALAPTATKYHA	188		
Db	147	YDLPATIDFIQKTRQBKHVHGHSQGTTGTFIAFTSNPALAKKIKRFYALPVAIVTKT	206		
Qy	189	KSPGTKEFLLPDMMIKGLFGKKEFL---YQTRFLRQLVIYLCGVVLDIQICSNMLLGG	245		
Db	207	ESPPKKIHFPKFKLVKIFVGNKMFPHNYLDQFLGTEV--CSRELLDLLCSNALFPCG	263		
Qy	246	PTNTNMNKSASVYAHTLAGTSVONILHVSQA VNSGELRAFDFMGSETKLEKNQPTFV	305		
Db	264	FDKKNLVNRDPDVLYGNRPAGTSGDILFHWAQLAKSKGLQYNWGSPLOMHLNQTTP	323		
Qy	306	RYRVDRMTVPMTMWGGQDWLSPNDPKMLSEVTNLIYHKNIKPEWAHVDFIGWLDAPH	365		

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2004, 14:14:34 ; Search time 32.1402 Seconds
(without alignments)
1926.040 Million cell updates/sec

Title: US-10-042-431-49

Perfect score: 2076

Sequence: 1 VHMPTKAVDPAPMNISEII.....IHLMQEETNLQSGRCRAVL 390

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2076	100.0	390	22	Human TANGO 294 ma
2	2076	100.0	409	22	Human CG162 (or C5
3	2076	100.0	423	22	Human TANGO 294.
4	2076	100.0	423	24	Amino acid sequenc
5	2005.5	96.6	423	23	Human lipid-associ
6	1991	95.9	398	23	Human lipid-lipase
7	1991	95.9	398	23	Human lipase prote
8	1536.5	74.0	371	23	Human lysosomal ac
9	1289	62.1	399	22	Human lysosomal ac

10	1289	62.1	399	22	Human lysosomal ac
11	1282	61.8	392	23	Human lysosomal ac
12	1181.5	56.9	449	22	Novel human secret
13	1165	56.1	221	22	Human TANGO 294 ex
14	1153	55.5	395	6	Rat lingual lipase
15	1150.5	55.4	398	14	RGL precursor. Or
16	1141.5	55.0	379	15	Canine gastric lip
17	1141.5	55.0	379	17	Dog gastric lipase
18	1141.5	55.0	380	15	Canine gastric lip
19	1133	54.6	395	23	Human triacylglyce
20	1133	54.6	399	23	Human lipid metabo
21	1129	54.3	399	22	Amino acid sequenc
22	1127	54.3	378	23	Kid goat ptegastr
23	1126	54.2	397	23	Human lysosomal ac
24	1109	53.4	398	7	Sequence of ptegas
25	1109	53.4	398	7	Sequence of human
26	1109	53.4	398	17	Human gastric lipa
27	1109	53.4	398	22	Human lipase prote
28	1109	53.4	398	24	Human gastric lipa
29	1093	52.6	403	23	Human lipid metabo
30	1082	52.1	427	23	Human lysosomal ac
31	1024	49.3	401	23	Human NOV6b protei
32	1023	49.3	395	23	Human lysosomal ac
33	1023	49.3	395	24	Human lipase prote
34	998	48.1	373	23	Human NOV1 protei
35	995.5	48.0	362	24	Human triacylglyce
36	969	46.7	365	23	Human lysosomal ac
37	909	43.8	331	24	Human DITHP secret
38	893	43.0	731	23	Human triacylglyce
39	890.5	42.9	390	23	Human NOV6a protei
40	783	37.7	144	22	Human TANGO 294 cy
41	767	36.9	276	23	Human triacylglyce
42	767	36.9	289	23	Human lipid metabo
43	731.5	35.2	280	23	Human triacylglyce
44	710.5	34.2	656	22	Novel human diagno
45	681	32.8	233	22	Human protein HP03

ALIGNMENTS

RESULT 1
AAB66067

ID AAB66067 standard; Protein; 390 AA.

AC AAB66067;

DT 30-MAR-2001 (first entry)

DE Human TANGO 294 mature protein.

KW TANGO protein; INTERCEPT protein; neurological disorder;

KW central nervous system; focal brain disorder; bipolar affective disorder;
KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
KW neuropsychiatric; psychoactive substance use; anxiety.

OS Homo sapiens.

PN WO2000077239-A2.

PD 21-DEC-2000.

PF 24-MAY-2000; 2000WO-US14858.

PR 14-JUN-1999; 99US-0333159.

PA (MILL-) MILLENNIUM PHARM INC.

PI McCarthy SA, Fraser CC, Sharp JD, Barnes TM;

DP WPI; 2001-032313/04.

XX TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
PT screening assays and diagnostic assays and for the treatment of
PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
PT disease -
XX
XX
PS Claim 8; Pages 324-325; 359pp; English.
XX
XX The present invention relates to TANGO or INTERCEPT proteins and coding
CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
CC sequences are useful for the treatment of neurological disorders such as
CC central nervous system (CNS) disorders, CNS-related disorders, focal
CC brain disorders, global-diffuse cerebral disorders and other
CC neurological and cerebrovascular disorders. The CNS disorders include
CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
CC autonomic function disorders such as hypertension and sleep disorders,
CC neuropsychiatric disorders, psychoactive substance use disorders,
CC anxiety, and bipolar affective disorder.
XX
XX Sequence 390 AA;

Query Match 100.0%; Score 2076; DB 22; Length 390;
Best Local Similarity 100.0%; Pred. No. 1.4e-204;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VHMPTKAVDPEAFMNISEIIHQGVPCBEYEVATEDGYILSVNRPGLVQPKTGRPV 60
Db 1 VHMPTKAVDPEAFMNISEIIHQGVPCBEYEVATEDGYILSVNRPGLVQPKTGRPV 60
QY 61 VLLQHLGVGASNWIISNLPNNISLGFILADAGFDVWVGNSRGNWSKHKHTLSIDQDFWA 120
Db 61 VLLQHLGVGASNWIISNLPNNISLGFILADAGFDVWVGNSRGNWSKHKHTLSIDQDFWA 120
QY 121 PSYDEMARFDPVAVINFLQKTQSEKIYVGVYSGTGTGFTAFSTWPELAQIKMYFALA 180
Db 121 PSYDEMARFDPVAVINFLQKTQSEKIYVGVYSGTGTGFTAFSTWPELAQIKMYFALA 180
QY 181 PIATVHAKSPGTFKLLPDMNKGFLGKKEFLYQTRFLQVLYLCCQVILDCQSNIM 240
Db 181 PIATVHAKSPGTFKLLPDMNKGFLGKKEFLYQTRFLQVLYLCCQVILDCQSNIM 240
QY 241 LLLGGNTNNMNSRVSVAHTLACTSVQNTLHWSQVNSGELRPFQWSETKNLEKN 300
Db 241 LLLGGNTNNMNSRVSVAHTLACTSVQNTLHWSQVNSGELRPFQWSETKNLEKN 300
QY 301 QPTPVRYRDMTPTAMWTGGQDMLSNPDEYKMLLSEVTNLIYHKNIPWAHVDFFIWL 360
Db 301 QPTPVRYRDMTPTAMWTGGQDMLSNPDEYKMLLSEVTNLIYHKNIPWAHVDFFIWL 360
QY 361 DAPHRMNEIHLMOQEETNLSQGRCEAVL 390
Db 361 DAPHRMNEIHLMOQEETNLSQGRCEAVL 390

RESULT 2
AAE11931
ID AAE11931 standard; Protein; 409 AA.
XX
AC AAE11931;
XX
XX 18-DEC-2001 (first entry)
XX
XX Human CG162 (or C59) lipase protein #2.
DE
XX Human; apolipoprotein; lipase; lipoprotein receptor; ALLr; angina;
KW cardiovascular disease; lipid metabolism; myocardial infarction;
KW cerebral ischaemia; arterial thrombosis; thrombolytic; antilipemic;
KW coronary artery thrombosis; cerebral artery thrombosis; stroke;
KW intracardiac thrombosis; gene therapy; cardiovascular; vasodilator;
KW neuroprotectant; cerebroprotective.

OS Homo sapiens.
XX Key Location/Qualifiers
FH Peptide 1..19 /label= Signal_peptide
FT Protein 20..409 /note= "Human mature CG162 (or C59) lipase protein"
XX
XX WO200179446-A2.
XX 25-OCT-2001.
XX 16-APR-2001; 2001WO-US12529.
XX 14-APR-2000; 2000US-197137P.
XX 20-JUN-2000; 2000US-0598042.
XX 03-AUG-2000; 2000US-0631451.
XX 22-SEP-2000; 2000US-0667298.
XX 17-NOV-2000; 2000US-0714936.
XX (HYSE-) HYSEQ INC.
XX
XX Ballinger DG, Loeb D, Montgomery JR, Tang TY, Zhou P, Goodrich R;
PI Liu C, Asundi V, Zhao QA, Wehrman T, Drmanac RT, Ren F, Qian XB;
PI Wang D;
XX
XX WPI; 2001-611724/70.
XX N-PSDB; AAD19226.
XX Nucleic acids encoding human apolipoproteins, lipases, and lipoprotein
PT receptor polypeptides, useful for preventing diagnosing and treating
PT lipid metabolism disorders, thrombosis and cardiovascular diseases -
XX
XX Claim 10; Fig 3; 266pp; English.

CC The invention relates to polynucleotides encoding proteins CG122, CG179,
CC CG95, CG121, CG162, CG27, CG153 and CG168 which are related to proteins
CC involved in lipid metabolism and cardiovascular disease such as human
CC apolipoproteins, lipases and lipoprotein receptor proteins. These DNA
CC and protein sequences are useful for treating or preventing disorders
CC associated with apolipoproteins, lipases and lipoprotein receptor (ALLr)
CC expression and for treating lipid metabolism, cardiovascular diseases
CC and thrombosis. Antibodies against these proteins are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of these sequences. ALLr polypeptides are also
CC useful for identifying agents (agonists and antagonists) that bind to
CC them and cells expressing ALLr proteins are useful for identifying a
CC therapeutic agent for use in treatment of a pathology related to
CC aberrant expression or physiological interactions of this polypeptide.
CC Vectors comprising these DNA and protein sequences are also useful for
CC producing ALLr proteins. The nucleic acids and polypeptides of the
CC invention are also useful for the treatment of occlusive cardiovascular
CC diseases, myocardial infarction, cerebral ischaemia, angina, arterial
CC thrombosis, coronary artery thrombosis and cerebral artery thrombosis
CC or intracardiac thrombosis and stroke. The nucleotides of the invention
CC are used in gene therapy. The present sequence is human CG162 (or C59)
CC lipase protein.

XX Sequence 409 AA;
QY Query Match 100.0%; Score 2076; DB 22; Length 409;
Best Local Similarity 100.0%; Pred. No. 1.5e-204;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VHMPTKAVDPEAFMNISEIIHQGVPCBEYEVATEDGYILSVNRPGLVQPKTGRPV 60
Db 20 VHMPTKAVDPEAFMNISEIIHQGVPCBEYEVATEDGYILSVNRPGLVQPKTGRPV 79
QY 61 VLLQHLGVGASNWIISNLPNNISLGFILADAGFDVWVGNSRGNWSKHKHTLSIDQDFWA 120
Db 80 VLLQHLGVGASNWIISNLPNNISLGFILADAGFDVWVGNSRGNWSKHKHTLSIDQDFWA 139
QY 121 PSYDEMARFDPVAVINFLQKTQSEKIYVGVYSGTGTGFTAFSTWPELAQIKMYFALA 180

Db 140 PSYDEWARFDPVAVINFLQKTGQEKIYVGYSGTMTGFIATPMTPELAQKIKYFALA 199
 QY 181 PIATVKHAKSPGTFKLLPDMIMKGLFGKKEFLYQTRFLRQLVLYLCGVILDOICSNIM 240
 Db 200 PIATVKHAKSPGTFKLLPDMIMKGLFGKKEFLYQTRFLRQLVLYLCGVILDOICSNIM 259
 QY 241 LLLGGFNTNNMMSRASVYAAHTLAGTSVQNIILHWSQAVNSGELRAFDPWGSSTKMLEKN 300
 Db 260 LLLGGFNTNNMMSRASVYAAHTLAGTSVQNIILHWSQAVNSGELRAFDPWGSSTKMLEKN 319
 QY 301 QPTPVRYRVDMTVPTAMWTGGQDWLSNPEDVYKMLLSEVTNLIYHKNIPEWAHVDFIWL 360
 Db 320 QPTPVRYRVDMTVPTAMWTGGQDWLSNPEDVYKMLLSEVTNLIYHKNIPEWAHVDFIWL 379
 QY 361 DAPHRMYNEIILHMQOEETNLSQGRCEAVL 390
 Db 380 DAPHRMYNEIILHMQOEETNLSQGRCEAVL 409

RESULT 3
 AAB66065
 ID AAB66065 standard; Protein; 423 AA.

XX AAB66065;

XX 30-MAR-2001 (first entry)

XX Human TANGO 294.

XX TANGO protein; INTERCEPT protein; neurological disorder;
 KW central nervous system; focal brain disorder; bipolar affective disorder;
 KW global diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
 KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
 KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
 KW neuropsychiatric; psychoactive substance use; anxiety.

XX Homo sapiens.

XX WO200077239-A2.

XX 21-DEC-2000.

XX 24-MAY-2000; 2000WO-US14858.

XX 14-JUN-1999; 99US-0333159.

XX (MILL-) MILLENNIUM PHARM INC.

XX McCarthy SA, Fraser CC, Sharp JD, Barnes TM;

XX WPI; 2001-032313/04.

XX N-PSDB; AAP45131, AAP45132.

XX TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
 PT screening assays and diagnostic assays and for the treatment of
 PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
 PT disease -

XX Claim 8; Fig 6; 359pp; English.

XX The present invention relates to TANGO or INTERCEPT proteins and coding
 CC sequences (see AAP45121-F45136 and AAP45138-F45139 and AAB66031-B66057,
 CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
 CC sequences are useful for the treatment of neurological disorders such as
 CC central nervous system (CNS) disorders, CNS-related disorders, focal
 CC brain disorders, global-diffuse cerebral disorders and other
 CC neurological and cerebrovascular disorders. The CNS disorders include
 CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
 CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
 CC autonomic function disorders such as hypertension and sleep disorders,
 CC neuropsychiatric disorders, psychoactive substance use disorders,
 CC anxiety, and bipolar affective disorder.

XX SQ Sequence 423 AA;
 Query Match 100.0%; Score 2076; DB 22; Length 423;
 Best Local Similarity 100.0%; Pred. No. 1.6e-204;
 Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VHMPTKAVDPEAFPMNISEIIHQGYPCPEYEVATEDGYILSVNRIPRGLVQPKKTGSRPV 60
 Db 34 VHMPTKAVDPEAFPMNISEIIHQGYPCPEYEVATEDGYILSVNRIPRGLVQPKKTGSRPV 93
 QY 61 VLLQHLGVGASGNWISNLPNNISLGFILADAGPDVVMGNSRGNANRKHKTLSIDQDEFWA 120
 Db 94 VLLQHLGVGASGNWISNLPNNISLGFILADAGPDVVMGNSRGNANRKHKTLSIDQDEFWA 153
 QY 121 PSYDEWARFDPVAVINFLQKTGQEKIYVGYSGTMTGFIATPMTPELAQKIKYFALA 180
 Db 154 PSYDEWARFDPVAVINFLQKTGQEKIYVGYSGTMTGFIATPMTPELAQKIKYFALA 213
 QY 181 PIATVKHAKSPGTFKLLPDMIMKGLFGKKEFLYQTRFLRQLVLYLCGVILDOICSNIM 240
 Db 214 PIATVKHAKSPGTFKLLPDMIMKGLFGKKEFLYQTRFLRQLVLYLCGVILDOICSNIM 273
 QY 241 LLLGGFNTNNMMSRASVYAAHTLAGTSVQNIILHWSQAVNSGELRAFDPWGSSTKMLEKN 300
 Db 274 LLLGGFNTNNMMSRASVYAAHTLAGTSVQNIILHWSQAVNSGELRAFDPWGSSTKMLEKN 333
 QY 301 QPTPVRYRVDMTVPTAMWTGGQDWLSNPEDVYKMLLSEVTNLIYHKNIPEWAHVDFIWL 360
 Db 334 QPTPVRYRVDMTVPTAMWTGGQDWLSNPEDVYKMLLSEVTNLIYHKNIPEWAHVDFIWL 393
 QY 361 DAPHRMYNEIILHMQOEETNLSQGRCEAVL 390
 Db 394 DAPHRMYNEIILHMQOEETNLSQGRCEAVL 423
 RESULT 4
 ABOU08369
 ID ABOU08369 standard; Protein; 423 AA.
 XX ABOU08369;
 DT 03-JUN-2003 (first entry)
 XX Amino acid sequence for human TANGO 294.
 KW Human; TANGO 294; INTERCEPT; cellular process; tissue typing;
 KW forensic biology; cellular protease activity; cell interaction;
 KW development; blood disorder; haematopoietic cell-related disorder;
 KW growth; cell proliferation; cell differentiation; gamma delta T-cell;
 KW immune system; metabolic disorder; homeostatic disorder; anaemia;
 KW developmental bone disorder; osteoporosis; bacterial infection;
 KW thrombocytopaenia; renal failure; liver disease; cystic fibrosis;
 KW obesity; cancer; bronchitis; asthma; emphysema; pulmonary oedema;
 KW respiratory distress syndrome; osteopathic; antibacterial; antianemic;
 KW thrombolytic; nephrotropic; antioesity; hepatotropic; cytostatic;
 KW antiinflammatory; antiasthmatic.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 XX Peptide 1..33
 FT /label= Predicted signal peptide
 FT /note= "Given as SEQ ID No:48 and specifically
 FT Claimed in Claim 8"
 FT Protein 34..423
 FT /label= Predicted mature TANGO_294 protein
 FT /note= "Given as SEQ ID No:49 and specifically
 FT Claimed in Claim 8"
 FT Domain 34..254
 FT /label= Extracellular domain
 FT /note= "Given as SEQ ID No:50 and specifically
 FT Claimed in Claim 8"

FT Domain 255..279 /label= Transmembrane domain /note= "Given as SEQ ID No:51 and specifically claimed in Claim 8"

FT Domain 280..423 /label= Cytoplasmic domain /note= "Given as SEQ ID No:52 and specifically claimed in Claim 8"

PN US2002182675-A1.

XX 05-DEC-2002.

XX 25-OCT-2001; 2001US-0042431.

XX 14-JUN-1999; 99US-0333159.

XX 24-MAY-2000; 2000US-0578063.

XX (MILL-) MILLENNIUM PHARM INC.

XX McCarthy SA, Fraser CC, Sharp JD, Barnes TW;

XX WPI; 2003-328617/31.

XX N-PSDB; ABX94106.

XX New TANGO and INTERCEPT proteins, useful as modulating agents in regulating a variety of cellular processes, in chromosome mapping, in tissue typing, and in forensic biology -

XX Claim 8; Fig 6A-6C; 232pp; English.

CC The present invention relates to the isolation of novel TANGO or INTERCEPT proteins, and the polynucleotide sequences encoding them.

CC The invention discloses sequences for human TANGO 202, TANGO 234, TANGO 265, TANGO 273, TANGO 286, TANGO 294 and INTERCEPT 296. The sequences for murine TANGO 202 and TANGO 273 are also provided. The TANGO polypeptides and the polynucleotide sequences encoding them are useful as modulating agents in regulating a variety of cellular processes. The polynucleotide sequences are useful as primers or hybridisation probes for the detection of nucleic acids encoding TANGO polypeptides, chromosome mapping, tissue typing, and in forensic biology. TANGO 202 can be used to diagnose, prevent or treat disorders relating to aberrant cellular protease activity, inappropriate interaction of cells with mediators, inappropriate development, and blood and haematopoietic cell-related disorders. TANGO 234 can be used to modulate growth, proliferation, survival, differentiation, and activity of gamma delta T-cells. TANGO 265 can be used to prevent, diagnose and treat disorders characterised by aberrant organisation or development of a tissue or organ, and for modulating differentiation of cells of the immune system. TANGO 273 is useful for diagnosing, treating or preventing e.g. metabolic, homeostatic and developmental bone disorders (e.g. osteoporosis), and bacterial infection. TANGO 286 is useful for treating or preventing e.g. anaemia, thrombocytopaenia, renal failure or liver disease. TANGO 294 is useful for treating or preventing e.g. cystic fibrosis or obesity. INTERCEPT 296 is useful for diagnosing, treating or preventing e.g. cancers, bronchitis, cystic fibrosis, asthma, emphysema, pulmonary oedema, or adult and infant respiratory distress syndromes. The present sequence represents human TANGO 294.

XX Sequence 423 AA;

Query Match Best Local Similarity 100.0%; Score 2076; DB 24; Length 423; Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VHMPTKAVDPKAFNISEIIHQYPCPEEYEVATEDGYILSVNRI PGLVQPKTGRPV 60

Db 34 VHMPTKAVDPKAFNISEIIHQYPCPEEYEVATEDGYILSVNRI PGLVQPKTGRPV 93

Qy 61 VLLQHLVGGASNWSNLPNNLSGLFADAGFDVWMCNSRGNWSRKHKTLSIDQDEFWA 120

Db 94 VLLQHLVGGASNWSNLPNNLSGLFADAGFDVWMCNSRGNWSRKHKTLSIDQDEFWA 153

Qy 121 PSYDEMARFDP LPAVINFILOKTQGEKIYYVGYSGTTMGFIATSTWPELAQKIMYFALA 180

Db 154 PSYDEMARFDP LPAVINFILOKTQGEKIYYVGYSGTTMGFIATSTWPELAQKIMYFALA 213

Qy 181 PIATVKAHSGPTKFLLLPDMMIKGLFGKEFLYQYRFLQYVLCGQVILQICSNIM 240

Db 214 PIATVKAHSGPTKFLLLPDMMIKGLFGKEFLYQYRFLQYVLCGQVILQICSNIM 273

Qy 241 LLLGFTNNNNMSRASVYAAHTLAGTSVONILHWSOAVNSGELRAPDWSGSEYKNEKCN 300

Db 274 LLLGFTNNNNMSRASVYAAHTLAGTSVONILHWSOAVNSGELRAPDWSGSEYKNEKCN 333

Qy 301 OPTPVYRVRDVTPTAMWTGGQDLNSDPEDVKMLSEVTNLIYHKNI PWAHVDFITWGL 360

Db 334 OPTPVYRVRDVTPTAMWTGGQDLNSDPEDVKMLSEVTNLIYHKNI PWAHVDFITWGL 393

Qy 361 DAPHRMNYNEIHLMQQEBETNLSQGRCEAVL 390

Db 394 DAPHRMNYNEIHLMQQEBETNLSQGRCEAVL 423

RESULT 5

ABG31894

ID ABG31894 standard; Protein; 423 AA.

XX AC ABG31894;

XX DT 05-NOV-2002 (first entry)

XX DE Human lipid-associated molecule, LIPAM.

XX Human; lipid-associated molecule; LIPAM; cancer; leukaemia; cardiovascular disorder; atherosclerosis; hypertension; stroke; angina pectoris; neurological disorder; Alzheimer's disease; epilepsy; dementia; Parkinson's disease; multiple sclerosis; anxiety; psoriasis; autoimmune disorder; inflammatory disorder; anaemia; asthma; trauma; acquired immunodeficiency syndrome; AIDS; bronchitis; diabetes mellitus; emphysema; Crohn's disease; atopic dermatitis; rheumatoid arthritis; gastrointestinal disorder; lipid metabolism disorder; transgenic animal.

OS Homo sapiens.

WO200246418-A2.

PD 13-JUN-2002.

XX 04-DEC-2001; 2001WO-US47430.

PR 08-DEC-2000; 2000US-254505P.

PR 15-DEC-2000; 2000US-256187P.

PR 22-DEC-2000; 2000US-257908P.

PR 26-JAN-2001; 2001US-264429P.

XX (INCY-) INCYTE GENOMICS INC.

XX Griffin JA, Gandhi AR, Ramkumar J, Tang YT, Ding L, Yue H;

PI Gietzen KJ, Sapperstein SK, Honchell CD, Bruns CM, Duggan BM;

PI Xu Y, Lee S;

XX WPI; 2002-619022/66.

DR N-PSDB; ABK90871.

XX Novel polypeptide of human lipid associated molecule, useful for diagnosing, treating and preventing cancer, hypercholesterolaemia, cirrhosis, myocardial infarction, Parkinson's disease, asthma, psoriasis, gastritis

PS Claim 1; Page 111-112; 126pp; English.

XX The invention relates to an isolated polypeptide (I) of human lipid-associated molecule (LIPAM), and an isolated polynucleotide (II) encoding

(I). (I), (II) and antibody to (II) are useful for treating a disease or condition associated with decreased or increased expression of functional LIPAM. (I) or (II) is useful for diagnosing, treating or preventing cancer e.g. leukemia, cancer of adrenal gland, bladder, bone, bone marrow, brain, ovary, etc; cardiovascular disorders e.g. atherosclerosis, hypertension, Raynaud's disease, angina pectoris, myocardial infarction, rheumatic fever, cardiomyopathy, pericarditis, chronic obstructive pulmonary disease, etc; neurological disorders e.g. Alzheimer's disease, stroke, epilepsy, dementia, Parkinson's disease, multiple sclerosis, Creutzfeldt-Jakob disease, myasthenia gravis, anxiety, diabetic neuropathy, etc; autoimmune/inflammatory disorders e.g. acquired immunodeficiency syndrome (AIDS), Addison's disease, anemia, asthma, bronchitis, diabetes mellitus, emphysema, Crohn's disease, atopic dermatitis, rheumatoid arthritis, psoriasis, systemic lupus erythematosus, trauma, etc; and gastrointestinal disorders e.g. gastritis, anorexia, nausea gastroenteritis, ulcerative colitis, cholecystitis, hepatitis, cirrhosis, hepatoma, diarrhoea, constipation, etc, and disorders of lipid metabolism e.g., fatty liver, cholestasis, Fabry's disease, hypercholesterolaemia. (I) or (II) is useful for assessing the effects of exogenous compounds on the expression of (I) or (II). (I) is useful in number of drug screening techniques, and to analyse the proteome of a tissue or cell type. (I) or antibody to (II) is useful as elements on a microarray. (II) is useful for creating knockin humanized animals or transgenic animals to model human diseases, in somatic or germline gene therapy, to generate a transcript image of a tissue or cell type, for detecting differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier or affected individuals, and as hybridisation probes for mapping naturally occurring genomic sequences. The present sequence represents the amino acid sequence of a human lipid-associated molecule (LIPAM) of the invention.

Sequence 423 AA;

Query Match 96.6%; Score 2005.5; DB 23; Length 423;
Best Local Similarity 97.2%; Pred. No. 2.8e-197;
Matches 379; Conservative 2; Mismatches 2; Indels 7; Gaps 1;
QY 8 VDPAFNISEIIHQGPCEEEVATGEGVILSVNRIPLGLVQPKTGRPVVLLQHL 67
Db 34 VNPEVMNTSEIIHQGPCEEEVATGEGVILSVNRIPLGLVQPKTGRPVVLLQHL 93
QY 68 VGGASNIISLNNLSGLFILLADAGFDVWNGSRGNWSRKHKTLSDQDEFWAFSYDEVA 127
Db 94 VGGASNIISLNNLSGLFILLADAGFDVWNGSRGNWSRKHKTLSDQDEFWAFSYDEVA 153
QY 128 RFDLPVAVINFLQKTGQEKIYYGYSGQTTMGFTAFSTMPPELAQKIMYPALAPIATVKH 197
Db 154 RFDLPVAVINFLQKTGQEKIYYGYSGQTTMGFTAFSTMPPELAQKIMYPALAPIATVKH 213
QY 188 AKSPGTFKLLLPDMWIKGLFGKKEFLYQTRFLRQVLYLGGVILDOICSNIMLLGGFN 247
Db 214 AKSPGTFKLLLPDMWIKGLFGKKEFLYQTRFLRQVLYLGGVILDOICSNIMLLGGFN 273
QY 248 TNNMNM-----SRASVYAAHTLAGTSVONILHWSQAVNSGELRAPDWGSETNLEKCN 300
Db 274 TNNMNMTHGLQSRASVYAAHTLAGTSVONILHWSQAVNSGELRAPDWGSETNLEKCN 333
QY 301 QPTPVRYVRDMVTPTAMWTGGQDLNPNEDVKMLSEVTNLIYHKNIPEWAHVDFIWL 360
Db 334 QPTPVRYVRDMVTPTAMWTGGQDLNPNEDVKMLSEVTNLIYHKNIPEWAHVDFIWL 393
QY 361 DAPHRMNEIHLMQQBEETNLSQGRCEAVL 390
Db 394 DAPHRMNEIHLMQQBEETNLSQGRCEAVL 423

RESULT 6
ABG71000
ID ABG71000 standard; Protein; 398 AA.
XX
AC ABG71000;
XX

DT 07-JAN-2003 (first entry)
XX Human lipase protein.
DE
XX Human; lipase; cancer; colon; kidney; skin; brain; testis; chromosome 10;
KW Burkitt's lymphoma; cytostatic; gene therapy; lipase-inhibitor.
XX Homo sapiens.
OS
XX Key Location/Qualifiers
PH Modified-site 88..91 /note= "N-glycosylation site"
FT Modified-site 260..263 /note= "N-glycosylation site"
FT Modified-site 388..391 /note= "N-glycosylation site"
FT Modified-site 22..25 /note= "cAMP- and cGMP-dependent protein kinase phosphorylation site"
FT Modified-site 113..115 /note= "Protein kinase C phosphorylation site"
FT Modified-site 192..194 /note= "Protein kinase C phosphorylation site"
FT Modified-site 130..133 /note= "Casein kinase II phosphorylation site"
FT Modified-site 150..153 /note= "Casein kinase II phosphorylation site"
FT Modified-site 173..176 /note= "Casein kinase II phosphorylation site"
FT Modified-site 335..338 /note= "Casein kinase II phosphorylation site"
FT Modified-site 149..157 /note= "Casein kinase II phosphorylation site"
FT Modified-site 74..79 /note= "Tyrosine kinase phosphorylation site"
FT Modified-site 77..82 /note= "N-myristoylation site"
FT Modified-site 105..110 /note= "N-myristoylation site"
FT Modified-site 109..114 /note= "N-myristoylation site"
FT Modified-site 163..168 /note= "N-myristoylation site"
FT Modified-site 252..257 /note= "N-myristoylation site"
FT Modified-site 215..218 /note= "N-myristoylation site"
FT Active-site 155..164 /note= "Amidation site"
FT /note= "Lipases, serine active site"
XX WO200274975-A2.
XX 26-SEP-2002.
XX 18-MAR-2002; 2002WO-US08035.
XX 20-MAR-2001; 2001US-0811825.
XX (PEKE) PE CORP NY.
XX Yan C, Di Francesco V, Beasley EM;
XX WPI; 2002-750560/81.
XX N-PSDB; ABS55356, ABS55357.
XX New isolated human lipase peptides and encoding nucleic acids, useful
PT for diagnosing and treating disorders mediated by human lipase
PT proteins, e.g. cancer of the colon, kidneys, skin and testis, and
PT Burkitt's lymphoma -
XX Claim 1; Fig 2; 79pp; English.
XX The present invention relates to a new lipase peptide. The methods and

compositions of the present invention are useful for diagnosing and treating disorders mediated by the human lipase protein, such as cancer of the colon, kidneys, skin, brain and testis, and Burkitt's lymphoma. The present amino acid sequence represents the human lipase protein of the invention. This sequence is encoded by the human lipase gene located on chromosome 10.

XX
SQ Sequence 398 AA;

Query Match 95.9%; Score 1991; DB 23; Length 398;
Best Local Similarity 100.0%; Pred. No. 8e-196;
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 SEIIHQGYPCBEYEYVATEDGYILSVNRIIPRGVLVQPKTGSFVLLQHLVGGASNNIS 76
DB 25 SEIIHQGYPCBEYEYVATEDGYILSVNRIIPRGVLVQPKTGSFVLLQHLVGGASNNIS 84

QY 77 NLPNNSLGFILADAGFDVWNGSRGNWAKSRKHTLSIDQDEFWAFSYDEMAFDFLPAVIN 136
DB 85 NLPNNSLGFILADAGFDVWNGSRGNWAKSRKHTLSIDQDEFWAFSYDEMAFDFLPAVIN 144

QY 137 FILQKTQGEKIYYVGYSGQTTMGFIAPSTMPPELAQKIKMYFALAPIATVGHAKSPGTFKL 196
DB 145 FILQKTQGEKIYYVGYSGQTTMGFIAPSTMPPELAQKIKMYFALAPIATVGHAKSPGTFKL 204

QY 197 LLPDMMIKGLFGKKEFLYQTRFLRQLVYLICQGVILDOICSNIMLLGGFNTNNMSRA 256
DB 205 LLPDMMIKGLFGKKEFLYQTRFLRQLVYLICQGVILDOICSNIMLLGGFNTNNMSRA 264

QY 257 SVYAAHTLAGTSVQNTILHWSQAVNSGELRAFDFWGSSTKLEKCNQTPVRYRVDMTVP 316
DB 265 SVYAAHTLAGTSVQNTILHWSQAVNSGELRAFDFWGSSTKLEKCNQTPVRYRVDMTVP 324

QY 317 AMWTGGQDWLSPEDVKMLLSEVTNLIYHKNIPFWAHVDFIWLGDAPHMYNEIHLMOQ 376
DB 325 AMWTGGQDWLSPEDVKMLLSEVTNLIYHKNIPFWAHVDFIWLGDAPHMYNEIHLMOQ 384

QY 377 EETNLSQGRCEAVL 390
DB 385 EETNLSQGRCEAVL 398

RESULT 7
AAU99164
ID AAU99164 standard; Protein; 398 AA.
AC AAU99164;
XX
DT 24-SEP-2002 (first entry)
XX
DE Human lysosomal acid lipase #1.
XX
KW Human; enzyme; lysosomal acid lipase; lipid malabsorption illness; cystic fibrosis; alcoholism; heart disease; heart attack;
KW Wolman disease; cholesterol ester storage disease; brain injury;
KW mood disorder; anxiety disorder; thought disorder; volition disorder;
KW sleep disorder; neurogenic disorder; myopathic disorder; COPD;
KW obesity; cancer; chronic obstructive pulmonary disease; diabetes;
KW cardiovascular disorder; Alzheimer's disease; Parkinson's disease;
KW anorexia; osteoarthritis; central nervous system disorder;
KW peripheral nervous system disorder.
XX
OS Homo sapiens.
XX
FN WO200236731-A2.
XX
PD 10-MAY-2002.
XX
PF 30-OCT-2001; 2001WO-BP12518.
XX
PR 31-OCT-2000; 2000US-244170P.
PR 29-MAY-2001; 2001US-293516P.
XX

(FARB) BAYER AG.
XX
FI Xiao Y;
XX
DR WPI; 2002-519248/55.
DR N-PSDB; ABK86569.
XX
PT Novel human lysosomal acid lipase polypeptide, useful for treating cancer, diabetes, obesity, chronic obstructive pulmonary disease, peripheral or central nervous system disorder of cardiovascular disorder

Claim 25; Fig 2; 126pp; English.

The invention relates to a purified human lysosomal acid lipase polypeptide. Also included are the polynucleotide encoding the lipase (or its fragment, derivative, allele or sequence at least 60% identical to it), vectors, host cells, a reagent (e.g. an antisense oligonucleotide) which binds to the lipase or polynucleotide (used for detection and modulating/reducing the lipase activity) and an anti-lipase antibody. The lipase and polynucleotide are useful for identifying therapeutic agents that either increase or decrease the lipase activity. The identified agent, the lipase and polynucleotide are useful for treatment of a disease such as lipid malabsorption illness, cystic fibrosis, alcoholism, heart disease, heart attack, Wolman disease, cholesterol ester storage disease, brain injury, mood disorder, anxiety disorder, thought disorder, volition disorder, sleep disorder, neurogenic disorder, myopathic disorder, obesity, cancer, chronic obstructive pulmonary disease (COPD), diabetes, cardiovascular disorder, Alzheimer's disease, Parkinson's disease, anorexia, osteoarthritis, a central nervous system disorder and a peripheral nervous system disorder. The present sequence is the human lysosomal lipase #1.

SQ Sequence 398 AA;

Query Match 95.9%; Score 1991; DB 23; Length 398;
Best Local Similarity 100.0%; Pred. No. 8e-196;
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 SEIIHQGYPCBEYEYVATEDGYILSVNRIIPRGVLVQPKTGSFVLLQHLVGGASNNIS 76
DB 25 SEIIHQGYPCBEYEYVATEDGYILSVNRIIPRGVLVQPKTGSFVLLQHLVGGASNNIS 84

QY 77 NLPNNSLGFILADAGFDVWNGSRGNWAKSRKHTLSIDQDEFWAFSYDEMAFDFLPAVIN 136
DB 85 NLPNNSLGFILADAGFDVWNGSRGNWAKSRKHTLSIDQDEFWAFSYDEMAFDFLPAVIN 144

QY 137 FILQKTQGEKIYYVGYSGQTTMGFIAPSTMPPELAQKIKMYFALAPIATVGHAKSPGTFKL 196
DB 145 FILQKTQGEKIYYVGYSGQTTMGFIAPSTMPPELAQKIKMYFALAPIATVGHAKSPGTFKL 204

QY 197 LLPDMMIKGLFGKKEFLYQTRFLRQLVYLICQGVILDOICSNIMLLGGFNTNNMSRA 256
DB 205 LLPDMMIKGLFGKKEFLYQTRFLRQLVYLICQGVILDOICSNIMLLGGFNTNNMSRA 264

QY 257 SVYAAHTLAGTSVQNTILHWSQAVNSGELRAFDFWGSSTKLEKCNQTPVRYRVDMTVP 316
DB 265 SVYAAHTLAGTSVQNTILHWSQAVNSGELRAFDFWGSSTKLEKCNQTPVRYRVDMTVP 324

QY 317 AMWTGGQDWLSPEDVKMLLSEVTNLIYHKNIPFWAHVDFIWLGDAPHMYNEIHLMOQ 376
DB 325 AMWTGGQDWLSPEDVKMLLSEVTNLIYHKNIPFWAHVDFIWLGDAPHMYNEIHLMOQ 384

QY 377 EETNLSQGRCEAVL 390
DB 385 EETNLSQGRCEAVL 398

RESULT 8
AAU99165
ID AAU99165 standard; Protein; 371 AA.
XX

AAU99165;
 24-SEP-2002 (first entry)
 Human lysosomal acid lipase #2.
 Human; enzyme; lysosomal acid lipase; lipid malabsorption illness;
 cystic fibrosis; alcoholism; heart disease; heart attack;
 Wolman disease; cholesterol ester storage disease; brain injury;
 mood disorder; anxiety disorder; thought disorder; volition disorder;
 sleep disorder; neurogenic disorder; myopathic disorder; COPD;
 obesity; cancer; chronic obstructive pulmonary disease; diabetes;
 cardiovascular disorder; Alzheimer's disease; Parkinson's disease;
 anorexia; osteoarthritis; central nervous system disorder;
 peripheral nervous system disorder.
 Homo sapiens.
 WO200236731-A2.
 10-MAY-2002.
 30-OCT-2001; 2001WO-EPI2518.
 31-OCT-2000; 2000US-244170P.
 29-MAY-2001; 2001US-293516P.
 (FARB) BAYER AG.
 Xiao Y;
 WPI; 2002-519248/55.
 N-PSDB; ABK6570.
 Novel human lysosomal acid lipase polypeptide, useful for treating
 cancer, diabetes, obesity, chronic obstructive pulmonary disease,
 peripheral or central nervous system disorder or cardiovascular
 disorder.
 Claim 25; Fig 5; 126pp; English.
 The invention relates to a purified human lysosomal acid lipase
 polypeptide. Also included are the polynucleotide encoding the
 lipase (or its fragment, derivative, allele or sequence at least 60%
 identical to it), vectors, host cells, a reagent (e.g. an antisense
 oligonucleotide) which binds to the lipase or polynucleotide (used for
 detection and modulating/reducing the lipase activity) and an anti-lipase
 antibody. The lipase and polynucleotide are useful for identifying
 therapeutic agents that either increase or decrease the lipase activity.
 The identified agent, the lipase and polynucleotide are useful for
 treatment of a disease such as lipid malabsorption illness,
 cystic fibrosis, alcoholism, heart disease, heart attack,
 Wolman disease, cholesterol ester storage disease, brain injury,
 mood disorder, anxiety disorder, thought disorder, volition disorder,
 sleep disorder, neurogenic disorder, myopathic disorder, obesity, cancer,
 chronic obstructive pulmonary disease (COPD), diabetes,
 cardiovascular disorder, Alzheimer's disease, Parkinson's disease,
 anorexia, osteoarthritis, a central nervous system disorder and
 a peripheral nervous system disorder. The present sequence is the
 human lysosomal lipase #2.

SQ Sequence 371 AA;

Query Match 74.0%; Score 1536.5; DB 23; Length 371;
 Best Local Similarity 80.9%; Pred. NO. 4.2e-149;
 Matches 297; Conservative 17; Mismatches 22; Indels 31; Gaps 2;

QY 17 SEIQHOGYPCBEVEVATEDGYILSVNRIPLGLVQPKTGSRPVLLQHLGVGASNWIS 76
 DB 25 SEIQHOGYPCBEVEVATEDGYILSVNRIPLGLVQPKTGSRPVLLQHLGVGASNWIS 84
 QY 77 NLPNNSLGFLILADAGFDVWMCNRSRGNWSRKHKTLSIDQDEFWAFSYDEMARDLPAVIN 136

DB 85 NLPNNSLGFLILADAGFDVWMCNRSRGNWSRKHKTLSIDQDEFWAFSYDEMARDLPAVIN 144
 QY 137 FILQKTGOEKIYYGYSGQTTMGFIAPSTWPELAKIKMYFALAPIATVGHAKSPGTFKL 196
 DB 145 FILQKTGOEKIYYGYSGQTTMGFIAPSTWPELAKIKMYFALAPIATVGHAKSPGTFKL 204
 QY 197 LLPDMWIKGLFGKKEFLYQTRFLRQLVYLCGQVILQICSNIMLLGGFNTNNMNM--- 253
 DB 205 LLPDMWIKGLFGKKEFLYQTRFLRQLVYLCGQVILQICSNIMLLGGFNTNNMNM 264
 QY 254 -----SRASYAAHTLAGTSVQNIILHWSQAVNSGELRAFDWGSFTKNLEKNCQPTPVYRV 309
 DB 265 GLLOSRSASYAAHTLAGTSVQNIILHWSQAVNSGELRAFDWGSFTKNLEKNCQPTPVYRV 300
 QY 310 RDMVTPTAMWTGGQDMLSNPDEVMALLSEVNTLYHKNIPDEWAVHVDFTWGLDAPHRMYNE 369
 DB 301 TKIEVPTAIWNGGQDIVADPKDVENLLPQIANLIYKLIPIHYNHVDVFLGSDAQEIYQD 360
 QY 370 IHLMOQ 376
 DB 361 LILMEE 367

RESULT 9

AA990783
 ID AAB90783 standard; Protein; 399 AA.

XX AAB90783;

XX 15-JUN-2001 (first entry)

DE Human shear stress-response protein SEQ ID NO: 66.

KW Human; shear stress-response protein; vascular disease;
 KW arteriosclerosis.

OS Homo sapiens.

PN WO200125427-A1.

PD 12-APR-2001.

PF 02-OCT-2000; 2000WO-JP06840.

PR 01-OCT-1999; 99JP-0280976.

XX (KIOW) KYOWA HAKKO KOGYO KK.

XX (NOUJ) NOUJIMA H.

XX Nojima H, Yoshigae H, Obayashi M, Ota T, Kawabata A, Sakurada K;
 PI Kuga T, Sekine S, Nakamura Y, Sugano S;

XX WPI; 2001-266308/27.

DR N-PSDB; AAH02906.

XX DNA sequences, proteins encoded by them and antibodies against them
 PT useful in diagnosis and treatment of vascular disease caused by
 PT arteriosclerosis.

PS Claim 60; Page 402-404; 678pp; Japanese.

CC The present invention provides the protein and coding sequences of a
 CC number of human shear stress response proteins. These are useful in the
 CC diagnosis, treatment and screening of vascular diseases caused by
 CC arteriosclerosis, including heart failure, post-PTCA restenosis and
 CC hypertension.

XX Sequence 399 AA;

Query Match 62.1%; Score 1289; DB 22; Length 399;
 Best Local Similarity 63.5%; Pred. NO. 1.4e-123;
 Matches 235; Conservative 55; Mismatches 80; Indels 0; Gaps 0;

QY 7 AVDPPEAFMNISEIIHQGYPCPEEYEVATEDGYILSVNRIPLGVLPKKTGSRPVVLLQHG 66
 DB 28 AVDPETNMVSEIIISYWGFPSEYLVETEDGYILCLNRIPLHGRKXHSKDGKPEVFLQHG 87
 QY 67 LVGGASNIWNLNPNNSLGFILADAGFDVVMGNSRGNAWSRKHKTLISIDODEFWAFSYDEM 126
 DB 88 LLADSSNWTNLANSLSLGFILADAGFDVVMGNSRGNTWSRKHKTLISVSDQDEFWAFSYDEM 147
 QY 127 ARFDLPVAVINFILOKTKQEKIYVYGVYSGTGMGFIATFSTMPPELAOKIKMYFALAPIATVK 186
 DB 148 AKYDLPAVINFILOKTKQEKIYVYGVYSGTGMGFIATFSTMPPELAOKIKMYFALAPIATVK 207
 QY 187 HAKSPGTFKLLPDMNKGFLGKKEFLYQTRFRLQRLVYLCQVILDOICSNIMLLGGF 246
 DB 208 FCTSPMAKLGRLPDHLIKDLFGDKFELPQSAFLKWLGTHTVTHILKELCGNLCFLLCGF 267
 QY 247 NTNNMNSRASVVAHTLACTSVQNLHWSQAVNSGELRAFQWGSSTKLEKNCNPTPVR 306
 DB 268 NERNLNMSRDVYTHSPAGTSVQNLHWSQAVNSGELRAFQWGSSTKLEKNCNPTPVR 327
 QY 307 YRVDMTVPTAMWTGGQDLNPNEDVYKMLLSEVNTLIYHKNIPPEWAHVDPIWGLDAPHRM 366
 DB 328 YNVKMDLVPVAVNSGGHDLADYVDVNNILLTQITLVFHSIPEWEHLDFIWLDPAPWRL 387
 QY 367 YNEIHLMOQ 376
 DB 388 YNKIINLMRK 397

RESULT 10
 AAB66061
 ID AAB66061 standard; Protein; 399 AA.
 AC AAB66061;
 XX
 DT 30-MAR-2001 (first entry)
 XX
 DE Human lysosomal acid lipase protein.
 XX
 KW Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
 KW central nervous system; focal brain disorder; bipolar affective disorder;
 KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
 KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
 KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
 KW neuropsychiatric; psychoactive substance use; anxiety.
 XX
 OS Homo sapiens.
 XX
 FN WO200077239-A2.
 XX
 PD 21-DEC-2000.
 XX
 PF 24-MAY-2000; 2000WO-US14858.
 XX
 PR 14-JUN-1999; 99US-0333159.
 XX
 FA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI McCarthy SA, Fraser CC, Sharp JD, Barnes TM;
 XX
 DR WPI; 2001-032313/04.
 XX
 TT TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
 PT screening assays and diagnostic assays and for the treatment of
 PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
 PT disease -
 XX
 PS Claim 8; Fig 6; 359pp; English.
 XX
 CC The present invention relates to TANGO or INTERCEPT proteins and coding
 CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
 CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
 CC sequences are useful for the treatment of neurological disorders such as

CC central nervous system (CNS) disorders, CNS-related disorders, focal
 CC brain disorders, global-diffuse cerebral disorders and other
 CC neurological and cerebrovascular disorders. The CNS disorders include
 CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
 CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
 CC autonomic function disorders, such as hypertension and sleep disorders,
 CC neuropsychiatric disorders, psychoactive substance use disorders,
 CC anxiety, and bipolar affective disorder. The present sequence is a
 CC sequence used in a sequence homology alignment with the TANGO/INTERCEPT
 CC sequences of the present invention.
 XX Sequence 399 AA;

QY 7 AVDPPEAFMNISEIIHQGYPCPEEYEVATEDGYILSVNRIPLGVLPKKTGSRPVVLLQHG 66
 DB 28 AVDPETNMVSEIIISYWGFPSEYLVETEDGYILCLNRIPLHGRKXHSKDGKPEVFLQHG 87
 QY 67 LVGGASNIWNLNPNNSLGFILADAGFDVVMGNSRGNAWSRKHKTLISIDODEFWAFSYDEM 126
 DB 88 LLADSSNWTNLANSLSLGFILADAGFDVVMGNSRGNTWSRKHKTLISVSDQDEFWAFSYDEM 147
 QY 127 ARFDLPVAVINFILOKTKQEKIYVYGVYSGTGMGFIATFSTMPPELAOKIKMYFALAPIATVK 186
 DB 148 AKYDLPAVINFILOKTKQEKIYVYGVYSGTGMGFIATFSTMPPELAOKIKMYFALAPIATVK 207
 QY 187 HAKSPGTFKLLPDMNKGFLGKKEFLYQTRFRLQRLVYLCQVILDOICSNIMLLGGF 246
 DB 208 FCTSPMAKLGRLPDHLIKDLFGDKFELPQSAFLKWLGTHTVTHILKELCGNLCFLLCGF 267
 QY 247 NTNNMNSRASVVAHTLACTSVQNLHWSQAVNSGELRAFQWGSSTKLEKNCNPTPVR 306
 DB 268 NERNLNMSRDVYTHSPAGTSVQNLHWSQAVNSGELRAFQWGSSTKLEKNCNPTPVR 327
 QY 307 YRVDMTVPTAMWTGGQDLNPNEDVYKMLLSEVNTLIYHKNIPPEWAHVDPIWGLDAPHRM 366
 DB 328 YNVKMDLVPVAVNSGGHDLADYVDVNNILLTQITLVFHSIPEWEHLDFIWLDPAPWRL 387
 QY 367 YNEIHLMOQ 376
 DB 388 YNKIINLMRK 397

RESULT 11
 AAO18227
 ID AAO18227 standard; Protein; 392 AA.
 XX
 AC AAO18227;
 XX
 DT 18-SEP-2002 (first entry)
 XX
 DE Human lysosomal acid lipase related protein.
 XX
 KW Human; lysosomal acid lipase; chromosome 10; lipemic; Wolman disease;
 KW cholesteryl ester storage disease.
 XX
 OS Homo sapiens.
 XX
 FN US6387680-B1.
 XX
 PD 14-MAY-2002.
 XX
 PF 29-MAR-2001; 2001US-0820001.
 XX
 PR 29-MAR-2001; 2001US-0820001.
 XX
 PI (PEKE) PE CORP NY.
 XX
 PI Merkuloov GV, Ketchum KA, Di Francesco V, Beasley EM;
 XX

XX DT 30-MAR-2001 (first entry)

XX DE Human TANGO 294 extracellular domain.

XX KW TANGO protein; INTERCEPT protein; neurological disorder;

XX KW central nervous system; focal brain disorder; bipolar affective disorder;

XX KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;

XX KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;

XX KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;

XX KW neuropsychiatric; psychoactive substance use; anxiety.

XX OS Homo sapiens.

XX PN W0200077239-A2.

XX PD 21-DEC-2000.

XX PF 24-MAY-2000; 2000MO-US14858.

XX PR 14-JUN-1999; 99US-0333159.

XX PA (MILL-) MILLENNIUM PHARM INC.

XX PI McCarthy SA, Fraser CC, Sharp JD, Barnes TM;

XX DR WPI; 2001-032313/04.

XX PT TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for

XX PT screening assays and diagnostic assays and for the treatment of

XX PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's

XX PT disease -

XX PS Claim 8; Page 326; 359pp; English.

XX CC The present invention relates to TANGO or INTERCEPT proteins and coding

XX CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,

XX CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding

XX CC sequences are useful for the treatment of neurological disorders such as

XX CC central nervous system (CNS) disorders, CNS-related disorders, focal

XX CC brain disorders, global-diffuse cerebral disorders and other

XX CC neurological and cerebrovascular disorders. The CNS disorders include

XX CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic

XX CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,

XX CC autonomic function disorders such as hypertension and sleep disorders,

XX CC neuropsychiatric disorders, psychoactive substance use disorders,

XX CC anxiety, and bipolar affective disorder.

XX SQ Sequence 221 AA;

Query Match 56.1%; Score 1165; DB 22; Length 221;

Best Local Similarity 100.0%; Pred. No. 3.1e-111;

Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VHMPTKAVDPAPFNNISEIIHQGPCEYEVEVATEDGYILSVNRIPRGLVQPKTKTSRPV 60

DB 1 VHMPTKAVDPAPFNNISEIIHQGPCEYEVEVATEDGYILSVNRIPRGLVQPKTKTSRPV 60

QY 61 VLLQGLVGGASNWSNLPNNLSGFTILADAGFDVWNGNSRGNWSRKHKTLSIDQDEFWA 120

DB 61 VLLQGLVGGASNWSNLPNNLSGFTILADAGFDVWNGNSRGNWSRKHKTLSIDQDEFWA 120

QY 121 PSYDEMAFDFLPAVINFLQTKGEBKYYVGYSGTMTGFIAPSTMPELAQIKMYFALA 180

DB 121 PSYDEMAFDFLPAVINFLQTKGEBKYYVGYSGTMTGFIAPSTMPELAQIKMYFALA 180

QY 181 PIATVGHAKSPGTFLLLPDMNMIKGLFGKKEFLYQTRFLRQ 221

DB 181 PIATVGHAKSPGTFLLLPDMNMIKGLFGKKEFLYQTRFLRQ 221

XX DT 30-MAR-2001 (first entry)

XX DE Human TANGO 294 extracellular domain.

XX KW TANGO protein; INTERCEPT protein; neurological disorder;

XX KW central nervous system; focal brain disorder; bipolar affective disorder;

XX KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;

XX KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;

XX KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;

XX KW neuropsychiatric; psychoactive substance use; anxiety.

XX OS Homo sapiens.

XX PN W0200077239-A2.

XX PD 21-DEC-2000.

XX PF 24-MAY-2000; 2000MO-US14858.

XX PR 14-JUN-1999; 99US-0333159.

XX PA (MILL-) MILLENNIUM PHARM INC.

XX PI McCarthy SA, Fraser CC, Sharp JD, Barnes TM;

XX DR WPI; 2001-032313/04.

XX PT TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for

XX PT screening assays and diagnostic assays and for the treatment of

XX PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's

XX PT disease -

XX PS Claim 8; Page 326; 359pp; English.

XX CC The present invention relates to TANGO or INTERCEPT proteins and coding

XX CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,

XX CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding

XX CC sequences are useful for the treatment of neurological disorders such as

XX CC central nervous system (CNS) disorders, CNS-related disorders, focal

XX CC brain disorders, global-diffuse cerebral disorders and other

XX CC neurological and cerebrovascular disorders. The CNS disorders include

XX CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic

XX CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,

XX CC autonomic function disorders such as hypertension and sleep disorders,

XX CC neuropsychiatric disorders, psychoactive substance use disorders,

XX CC anxiety, and bipolar affective disorder.

XX SQ Sequence 221 AA;

Query Match 56.1%; Score 1165; DB 22; Length 221;

Best Local Similarity 100.0%; Pred. No. 3.1e-111;

Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VHMPTKAVDPAPFNNISEIIHQGPCEYEVEVATEDGYILSVNRIPRGLVQPKTKTSRPV 60

DB 1 VHMPTKAVDPAPFNNISEIIHQGPCEYEVEVATEDGYILSVNRIPRGLVQPKTKTSRPV 60

QY 61 VLLQGLVGGASNWSNLPNNLSGFTILADAGFDVWNGNSRGNWSRKHKTLSIDQDEFWA 120

DB 61 VLLQGLVGGASNWSNLPNNLSGFTILADAGFDVWNGNSRGNWSRKHKTLSIDQDEFWA 120

QY 121 PSYDEMAFDFLPAVINFLQTKGEBKYYVGYSGTMTGFIAPSTMPELAQIKMYFALA 180

DB 121 PSYDEMAFDFLPAVINFLQTKGEBKYYVGYSGTMTGFIAPSTMPELAQIKMYFALA 180

QY 181 PIATVGHAKSPGTFLLLPDMNMIKGLFGKKEFLYQTRFLRQ 221

DB 181 PIATVGHAKSPGTFLLLPDMNMIKGLFGKKEFLYQTRFLRQ 221

RESULT 15
AAR37302
ID AAR37302 standard; Protein; 398 AA.
XX AC AAR37302;
XX DT 25-MAR-2003 (updated)
XX DT 20-SEP-1993 (first entry)
XX DE RGL precursor.
XX KW Rabbit gastric lipase; RGL; pRGLM2.1; fat; bioconversion;
XX KW hydrolysis; transesterification.
XX OS Oryctolagus cuniculus.
FH Key Location/Qualifiers
FT Protein 23..398
/note= "claim 1; page 10-11"
XX EP542629-A1.
XX PD 19-MAY-1993.
XX PF 12-NOV-1992; 92EP-0403055.
XX PR 13-NOV-1991; 91FR-0013948.
XX PA (LJOU) INST RECH JOUVEINAL.
XX PI Benicourt C, Blanchard C, Junien J;
XX DR WPI; 1993-161080/20.
XX DR N-PSDB; AAQ42310.
XX PT Rabbit gastric lipase, its precursor and their DNA - useful for
XX PT treating conditions linked to gastric lipase deficiency, such as
XX PT mucoviscidiosis and pancreatic exocrine insufficiency
PS Claim 1; Fig 7; 31pp; French.
XX RGL, opt. used with other lipases, are useful therapeutically
XX CC (1) to facilitate absorption of ingested fats in patients deficient
XX CC in endogenous gastric lipase and (2) to treat disorders caused by
XX CC inadequate lipase prodn. esp. mucoviscidiosis or pancreatic exocrine
XX CC deficiency. It can also be used (partic. when immobilised) for
XX CC enzymatic bioconversion, e.g. hydrolysis or transesterification.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
SQ Sequence 398 AA;

Query Match 55.4%; Score 1150.5; DB 14; Length 398;
Best Local Similarity 56.8%; Pred. No. 2.4e-109;
Matches 213; Conservative 65; Mismatches 94; Indels 3; Gaps 1;

QY 4 PTKAVDPEAFNMBISIIHQGQPCBEYEYATEDGYILSVNRIPRGLVQPKTKGSRPVLL 63
DB 26 PT---NPEVNNNISQMSISYWGYPSEKYEVTEDGYILEVNRIPYKKNNGRQRPVFL 82
QY 64 QHGLVGGASNWIENLPNNSLGFIADAGFDVVMGNSRGNWSRKHKTLISIDQDEWAFSY 123
DB 83 QHGLLASASNWIENLPNNSLAFILADAGYVGLGNSRGNWTSRNLNLYSPDSVEFWAFSF 142
QY 124 DEMARFDLPVNFITLOKTGQKIVYGVSGTGTFTARSTMPDLAKIKMYFALAPIA 183
DB 143 DEMAKYDLPATIDFIVKGTGQKIVYGVSGTGTFTARSTMPDLAKIKMYFALAPIA 202
QY 184 TVKHAKSGTGFLLLPDMMIKGLFGKKEFLYQTRFLRQLYLCCGVILQICSNIMLL 243
DB 203 TVKYTKSLVNLKRFIPPTMFKIIFGDKIFYPHNPFDDQFLATQVCSRETNVICSNALFII 262
QY 244 GGFNTNNNMRSASYAAHTLAGTSVQNLHWSQAVNSGELRAFWDGSETKLEKCNQPT 303

Search completed: February 19, 2004, 14:20:38
Job time : 34.1402 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2004, 14:19:24 ; Search time 13.4317 Seconds
(without alignments)
1228.526 Million cell updates/sec

Title: US-10-042-431-49

Perfect score: 2076

Sequence: 1 VHMPTKAVDPPEAFNMISEII.....IHLMQEETNLSQRCBAVL 390

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgm2_6/ptodata/2/iaa/5A.COMB.pep.*
- 2: /cgm2_6/ptodata/2/iaa/5B.COMB.pep.*
- 3: /cgm2_6/ptodata/2/iaa/6A.COMB.pep.*
- 4: /cgm2_6/ptodata/2/iaa/6B.COMB.pep.*
- 5: /cgm2_6/ptodata/2/iaa/PTCUS.COMB.pep.*
- 6: /cgm2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1282	61.8	392	4	US-09-820-001-4
2	1161	55.9	377	1	US-08-227-108-17
3	1161	55.9	377	2	US-09-073-674-17
4	1150.5	55.4	379	1	US-08-227-108-18
5	1150.5	55.4	379	2	US-09-073-674-18
6	1141.5	55.0	375	4	US-09-348-930A-6
7	1141.5	55.0	379	1	US-08-227-108-3
8	1141.5	55.0	379	2	US-09-073-674-3
9	1141.5	55.0	379	4	US-09-348-930A-2
10	1141.5	55.0	380	1	US-08-227-108-5
11	1141.5	55.0	380	2	US-09-073-674-5
12	1127	54.3	378	4	US-09-186-489-2
13	1127	54.3	378	4	US-10-043-665B-2
14	1109	53.4	398	4	US-09-348-930A-9
15	1098.5	52.9	380	1	US-08-227-108-16
16	1098.5	52.9	380	2	US-09-073-674-16
17	1023	49.3	395	4	US-09-820-001-2
18	1003	48.3	325	4	US-09-348-930A-4
19	143.5	6.9	348	4	US-09-252-991A-30091
20	105	5.1	430	1	US-08-484-105-22
21	105	5.1	430	1	US-08-484-106-22
22	103	5.0	292	4	US-09-328-352-7538
23	98.5	4.7	388	1	US-08-232-519-2
24	98.5	4.7	388	1	US-08-456-956-2
25	96.5	4.6	287	4	US-09-584-568C-4
26	96.5	4.6	1752	4	US-09-865-621A-2
27	95	4.6	346	2	US-08-602-359A-34

28	91	4.4	298	4	US-09-355-166-6	Sequence 6, Appli
29	90	4.3	277	4	US-09-424-349A-6	Sequence 6, Appli
30	89	4.3	312	4	US-09-107-532A-5684	Sequence 5684, Ap
31	89	4.3	494	4	US-09-134-001C-4475	Sequence 4475, Ap
32	89	4.3	684	4	US-09-252-991A-18416	Sequence 18416, A
33	89	4.3	1674	2	US-08-968-542C-12	Sequence 12, Appl
34	88	4.2	826	4	US-09-328-352-7515	Sequence 7515, Ap
35	87.5	4.2	297	4	US-09-584-568C-7	Sequence 7, Appli
36	87	4.2	270	4	US-09-107-532A-4778	Sequence 4778, Ap
37	87	4.2	624	2	US-08-756-317-9	Sequence 9, Appli
38	86.5	4.2	685	4	US-08-671-757A-12	Sequence 12, Appl
39	86.5	4.2	869	2	US-08-483-101-15	Sequence 15, Appl
40	84.5	4.1	173	4	US-09-328-352-8078	Sequence 8078, Ap
41	84.5	4.1	279	4	US-09-328-352-5958	Sequence 5958, Ap
42	84.5	4.1	338	2	US-08-602-359A-40	Sequence 40, Appl
43	84	4.0	347	4	US-09-328-352-5921	Sequence 5921, Ap
44	84	4.0	719	3	US-09-386-607-2	Sequence 2, Appli
45	84	4.0	719	4	US-09-645-707B-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-820-001-4
; Sequence 4, Application US/09820001
; Patent No. 6387680
; GENERAL INFORMATION:
; APPLICANT: MERKULOV, Gennady et al
; TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001186
; CURRENT APPLICATION NUMBER: US/09/820,001
; CURRENT FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Human
US-09-820-001-4

Query Match	61.8%	Score 1282;	DB 4;	Length 392;
Best Local Similarity	63.8%	Pred. No. 3.1e-129;		
Matches	234;	Conservative	53;	Mismatches 80;
		Indels	0;	Gaps 0;
Qy	7	AVDPEAFNMISEIIHQGYPCPEEYEVATEDGYILSVNRIPLGLVQPKTGRPVVLLQHG	66	
Db	26	AVDPTNMNVSEIIITWGPFPSEIYVETEDGYILCNRIPLGRKNHSKGPFPVFLQHG	85	
Qy	67	LVGGASNWTISNLPNNSLGFILADAGFDVVMGNSRGNWSRKHKTLSIDQDEFWAFSYDEM	126	
Db	86	LLADSSNVTNLSANGSLGFILADAGFDVVMGNSRGNWSRKHKTLSVSDQDEFWAFSYDEM	145	
Qy	127	AFDIPAVINFILOKTQGEKIYVGYSGTTFGTAFSTMTPELAQKIMYFALAPIATVK	186	
Db	146	AKYDLPASINFILKTKGQEVYVGHSGQTITGTFATFQIPELAKRIKMFALGPVAVSA	205	
Qy	187	HAKSPGTFKLLPLDMWIKGLFGKKEFLYQTRFLRLQVLYLQGVILDOQICSNIMLLQGF	246	
Db	206	FCTSPMAKLGRLPDHLIKDLFGDKFELPQSAFLKWLGTHTVILKELCGNLCLQGF	265	
Qy	247	NTNNMNSASVYAHTLAGTSVQNLHWSQVNSGELRAPDWSGSETNLEKCNQTPVR	306	
Db	266	NEERLNMNSVDYVTHTHSPAGTSVQNLHWSQVNFQKQFQDFWGSASAKNYHYNSYPT	325	
Qy	307	YVRDNTVPTAMWTGQDMLSNPDEVKMLLSVNTLIYHKNIPEWAHVDPIWGLDAPHRM	366	
Db	326	YVCKMLVPTAVWSGGHDLADVDVNIILITQITNLVPHESIPWEHLDPIWGLDAPWEL	385	
Qy	367	YNEIHL	373	

Db 386 YNKIINL 392

RESULT 2

US-08-227-108-17
; Sequence 17, Application US/08227108
; Patent No. 5807726
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/227,108
; FILING DATE: 03-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fanucci, Allan A.
; REGISTRATION NUMBER: 30,256
; REFERENCE/DOCKET NUMBER: 7620-033
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 377 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-227-108-17

Query Match 55.9%; Score 1161; DB 1; Length 377;
Best Local Similarity 57.7%; Pred. No. 3e-116;
Matches 213; Conservative 63; Mismatches 93; Indels 0; Gaps 0;
Qy 9 DPEAFMNISETIIOHQYPCBEYEVATEDGYILSVNRIIPRLGVOPKKTGSRPVVLLQGLV 68
Db 9 NPEANNNISQMITTWGYPCEYEVVTDGYILGVYRIPHGKNNSENIGKRPVVYLQGLI 68
Qy 69 GGASNWTLSNPNNSLGFILADAGFDVWNGSRGNAMSRKHKTLSIDODEFWAFSPYDEMAR 128
Db 69 ASATNWTIANLPNSLAFMLADAGDYVWLGNSRGNTWSRKVYSPDSVEFWAFSPDEMAK 128
Qy 129 FDLPAVINFILOKQTSKIYVGVYSGQTTMGFTAFSTMPPELAQIKMYFALAPIATVYKHA 188
Db 129 YDLPATINFILOKQTSKIYVGVYSGQTTMGFTAFSTMPPELAQIKMYFALAPIATVYKHA 188
Qy 189 KSPGTKEFLLLPDMNKGFLGKCEFLYOTRFLQRLVYLCQVILDOICSNIMLLGGENT 248
Db 189 QSLPKLISFTFTFLKLMFGKMFPHYTFDDFLGTGTEVCSREVLDDLLCSNTLFIKCFGDK 248
Qy 249 NNAMNSRASYAAHTLAGTSVQNTLHWSQAVNSGELRAFDPWGETKNEKCNQPTPVYR 308
Db 249 KNLNVSFVYLGHPAGTSVQDFLHWAQLVRSKGFQAFNWSGFSQMLHYNKQTPPEYD 308
Qy 309 VRDMTVPTAMKTCQDMLSNPEVYKMLLSVNTNLIYHKNIPENAHVDFIWLGLDAPHRMYN 368
Db 309 VSAITVPAVWNGNDILADPDQVAMLLPKLSNLLFHKELIAYNHLDFIWMADPAQEVYN 368

Qy 369 EIIHMQOE 377
Db 369 EMISMAED 377

RESULT 3

US-09-073-674-17
; Sequence 17, Application US/09073674
; Patent No. 5998189
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Warner-Lambert Company
; STREET: 2800 Plymouth Road
; CITY: Ann Arbor
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 48105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,674
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Crissey, Todd M.
; REGISTRATION NUMBER: 37,807
; REFERENCE/DOCKET NUMBER: 5072-D1-66-TMC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 734 622-7530
; TELEFAX: 734 622-1553
; TELEX:
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 377 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-073-674-17

Query Match 55.9%; Score 1161; DB 2; Length 377;
Best Local Similarity 57.7%; Pred. No. 3e-116;
Matches 213; Conservative 63; Mismatches 93; Indels 0; Gaps 0;
Qy 9 DPEAFMNISETIIOHQYPCBEYEVATEDGYILSVNRIIPRLGVOPKKTGSRPVVLLQGLV 68
Db 9 NPEANNNISQMITTWGYPCEYEVVTDGYILGVYRIPHGKNNSENIGKRPVVYLQGLI 68
Qy 69 GGASNWTLSNPNNSLGFILADAGFDVWNGSRGNAMSRKHKTLSIDODEFWAFSPYDEMAR 128
Db 69 ASATNWTIANLPNSLAFMLADAGDYVWLGNSRGNTWSRKVYSPDSVEFWAFSPDEMAK 128
Qy 129 FDLPAVINFILOKQTSKIYVGVYSGQTTMGFTAFSTMPPELAQIKMYFALAPIATVYKHA 188
Db 129 YDLPATINFILOKQTSKIYVGVYSGQTTMGFTAFSTMPPELAQIKMYFALAPIATVYKHA 188
Qy 189 KSPGTKEFLLLPDMNKGFLGKCEFLYOTRFLQRLVYLCQVILDOICSNIMLLGGENT 248
Db 189 QSLPKLISFTFTFLKLMFGKMFPHYTFDDFLGTGTEVCSREVLDDLLCSNTLFIKCFGDK 248
Qy 249 NNAMNSRASYAAHTLAGTSVQNTLHWSQAVNSGELRAFDPWGETKNEKCNQPTPVYR 308
Db 249 KNLNVSFVYLGHPAGTSVQDFLHWAQLVRSKGFQAFNWSGFSQMLHYNKQTPPEYD 308

QY 309 VRDVTPTAMTGGQDWSNPEDYKMLLSEVTNLIYHKNIPDEWAVHDFIWLGLDAPHRMYN 368
Db 309 VSAMTVPVAVWNGNDIILADPDQVAMLLFLKLSLLFHKELIAYNLHDFIWMADAPQEVYN 368
QY 369 EIIHLMQOE 377
Db 369 EMISMAED 377

RESULT 4
US-08-227-108-18
; Sequence 18, Application US/08227108
; Patent No. 5807726
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/227.108
; FILING DATE: 03-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Panucci, Allan A.
; REGISTRATION NUMBER: 30,256
; REFERENCE/DOCKET NUMBER: 7620-033
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-227-108-18

Query Match 55.4%; Score 1150.5; DB 1; Length 379;
Best Local Similarity 56.8%; Pred. No. 4.1e-115;
Matches 213; Conservative 65; Mismatches 94; Indels 3; Gaps 1;

QY 4 PTKAVDPEAFMNISEIIHQHGPCYEYEVATEDGVILSVNRIPLGVQPKTGRSPVLL 63
Db 7 PT---NPEVNMNISQMSIYWGYPSEKYEYVETEDGVILEVNRIPYKGNKSGNRGQRPVFL 63
QY 64 QHGLVGASNWTLSNLPNNSLGFLADAGFDVVMGNSRGNWSRKHKTLSDIDQDEFWAFSY 123
Db 64 QHGLLASASNWTLSNLPNNSLAFILADAGYGVWLGNSRGTWNSRRNLYYSPDSVEFWAFSP 123
QY 124 DEMARFDLPVINFILQKTQGEKIYVYVGSQGTTFGFIAPSTMPPELAQIKMYFALAPIA 183
Db 124 DEMAKYDLPATIDFIVKETQGEKLYHVGHSQGTTFGFIAPSTNPKLAERIKTFYALAPVA 183
QY 184 TVKHAKSPQTKFLLPDMNMIKGLFGKGFYQTRFLRQLVLYLCQVILDOICSNIMLL 243
Db 184 TVKYTKSLVNKLRFPPTMFKIIFGDKIPYHPNFPDQFLATQVCSRETLNVCNALLFI 243
QY 244 GGFNTNNMWSRASYAAHTLAGTSVQNLHWSQVNSGELRAFDWGSSETKQLEKCNQPT 303

Db 244 CQFDSANLNSRLDVTSHNPAGTSVQNLHWTQVKGSGNFOAFNGSPAQNVVHFNOPT 303
QY 304 PVRVRYRDVTPTAMTGGQDWSNPEDYKMLLSEVTNLIYHKNIPDEWAVHDFIWLGLDAP 363
Db 304 PPTVNTAMNVEIAVWNGNDIILADPDQVAMLLFLKLSLLFHKELIAYNLHDFIWMADAP 363
QY 364 HMYNEIHLMOOE 378
Db 364 QEVINEIISMMARK 378

RESULT 5
US-09-073-674-18
; Sequence 18, Application US/09073674
; Patent No. 5998189
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Warner-Lambert Company
; STREET: 2800 Plymouth Road
; CITY: Ann Arbor
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 48105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,674
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Criesey, Todd M.
; REGISTRATION NUMBER: 37,807
; REFERENCE/DOCKET NUMBER: 5072-DI-66-TWC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 734 622-7530
; TELEFAX: 734 622-1553
; TELEX:
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-073-674-18

Query Match 55.4%; Score 1150.5; DB 2; Length 379;
Best Local Similarity 56.8%; Pred. No. 4.1e-115;
Matches 213; Conservative 65; Mismatches 94; Indels 3; Gaps 1;

QY 4 PTKAVDPEAFMNISEIIHQHGPCYEYEVATEDGVILSVNRIPLGVQPKTGRSPVLL 63
Db 7 PT---NPEVNMNISQMSIYWGYPSEKYEYVETEDGVILEVNRIPYKGNKSGNRGQRPVFL 63
QY 64 QHGLVGASNWTLSNLPNNSLGFLADAGFDVVMGNSRGNWSRKHKTLSDIDQDEFWAFSY 123
Db 64 QHGLLASASNWTLSNLPNNSLAFILADAGYGVWLGNSRGTWNSRRNLYYSPDSVEFWAFSP 123
QY 124 DEMARFDLPVINFILQKTQGEKIYVYVGSQGTTFGFIAPSTMPPELAQIKMYFALAPIA 183
Db 124 DEMAKYDLPATIDFIVKETQGEKLYHVGHSQGTTFGFIAPSTNPKLAERIKTFYALAPVA 183
QY 184 TVKHAKSPQTKFLLPDMNMIKGLFGKGFYQTRFLRQLVLYLCQVILDOICSNIMLL 243

Db 184 TVKTKSLVNLKRLPIPTMTFKIIFGDKIFYPHNFFDQFLATQVCSRETNLVICSNALFTI 243

Qy 244 GGFNTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNSGELAFDWSGSETKNLEKCNQPT 303

Db 244 CGFDSANLNSRLDVTYVSHNPAGTSVQNLHWSQAVNSGELAFDWSGSETKNLEKCNQPT 303

Qy 304 PVRVYRDMVTPTAMTGGODWLSNPDVWMLSEVNTLYHKNIPFWAHVDFIWLGLDAP 363

Db 304 PPTVNTAMNVPITAVWSGGNDLADPDVLLPKULSNLYHKILPYNHLDPIFMANAP 363

Qy 364 HRMYNEIHLMOQEE 378

Db 364 QEVINEIISMAXDK 378

RESULT 6

US-09-348-930A-6

Sequence 6, Application US/09348930A

Patent No. 6573431

GENERAL INFORMATION:

APPLICANT: Lenee, P

APPLICANT: Gruber, V

APPLICANT: Baudino, S

APPLICANT: Merot, B

APPLICANT: Benicourt, C

APPLICANT: Cudrey, C

TITLE OF INVENTION: Recombinant Pseudocatal Lipases and Polypeptides Derivatives Produced Therefrom

TITLE OF INVENTION: Plants, Processes for Obtaining Them and Their Uses

FILE REFERENCE: 18433/2012

CURRENT APPLICATION NUMBER: US/09/348,930A

CURRENT FILING DATE: 1999-07-02

PRIOR APPLICATION NUMBER: 08/945,321

PRIOR FILING DATE: 1998-02-12

PRIOR APPLICATION NUMBER: PCT/FR96/00606

PRIOR FILING DATE: 1996-03-19

NUMBER OF SEQ ID NOS: 37

SOFTWARE: Patent in version 3.0

SEQ ID NO 6

LENGTH: 375

TYPE: PRT

ORGANISM: Canis familiaris

US-09-348-930A-6

Query Match 55.0%; Score 1141.5; DB 4; Length 375;

Best Local Similarity 56.5%; Pred. No. 3.7e-114;

Matches 210; Conservative 67; Mismatches 90; Indels 5; Gaps 3;

Qy 4 PTKAVDPEAFNISEIIOHQGYPCPEEVEVATEDGYILSVNRIPLGLVQPKTKGSRPVLL 63

Db 3 PT---NPEVTNISQMITYNGYPAEVEVATEDGYILGIDRIPIYGRKNSENIGRPPVAF 59

Qy 64 QHGLVGGASNWSNLNNLSGFIILADAGFDVWNGNSRGNWSRKHKTLSIDQDEFWAFSY 123

Db 60 QHGLASATNWSNLNNLSGFIILADAGFDVWNGNSRGNWSRKHKTLSIDQDEFWAFSF 119

Qy 124 DEMARFDLPVINEITLOKTQGEKIYYGYSGTMTGFIAPSTWPELAOKIKMYFALAPIA 183

Db 120 DEMAKYDLPATIDFILLKTKGQDKLHYVHSGTITGFIAPSTWPELAOKIKMYFALAP 179

Qy 184 TVKHAQSPGTFKLLPDMNMIKGLFGKKEFLYQTRFLRQ-LVIVLCGGVILDOICSNIMLL 242

Db 180 TVKYTETLLNKLMLVPSFLPKLIFGNKIF-YPHHFFDQFLATEVCSRETVDLLCSNALFI 238

Qy 243 LGGFNTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNSGELAFDWSGSETKNLEKCNQPT 302

Db 239 ICGFDTNNLNSRLDVTYVSHNPAGTSVQNLHWSQAVNSGELAFDWSGSETKNLEKCNQPT 302

Qy 303 TPVRYRDMVTPTAMTGGODWLSNPDVWMLSEVNTLYHKNIPFWAHVDFIWLGLDAP 362

Db 299 MPPTVNTLDMHVPITAVWSGGNDLADPDVLLPKULSNLYHKILPYNHLDPIFMANAP 358

Qy 363 PHRYNEIHLMOQEE 378

Db 359 POAVNEIIVSM 370

RESULT 7

US-08-227-108-3

Sequence 3, Application US/08227108

Patent No. 5807726

GENERAL INFORMATION:

APPLICANT: Blanchard, Claire

APPLICANT: Benicourt, Claude

APPLICANT: Junien, Jean-Louis

TITLE OF INVENTION: Recombinant Dog Gastric Lipase

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/227,108

FILING DATE: 03-APR-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Fanucci, Allan A.

REGISTRATION NUMBER: 30,256

REFERENCE/DOCKET NUMBER: 7620-033

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212 790-9090

TELEFAX: 212 869-8864/9741

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 379 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-227-108-3

Query Match 55.0%; Score 1141.5; DB 1; Length 379;

Best Local Similarity 56.5%; Pred. No. 3.8e-114;

Matches 210; Conservative 67; Mismatches 90; Indels 5; Gaps 3;

Qy 4 PTKAVDPEAFNISEIIOHQGYPCPEEVEVATEDGYILSVNRIPLGLVQPKTKGSRPVLL 63

Db 7 PT---NPEVTNISQMITYNGYPAEVEVATEDGYILGIDRIPIYGRKNSENIGRPPVAF 63

Qy 64 QHGLVGGASNWSNLNNLSGFIILADAGFDVWNGNSRGNWSRKHKTLSIDQDEFWAFSY 123

Db 64 QHGLASATNWSNLNNLSGFIILADAGFDVWNGNSRGNWSRKHKTLSIDQDEFWAFSF 123

Qy 124 DEMARFDLPVINEITLOKTQGEKIYYGYSGTMTGFIAPSTWPELAOKIKMYFALAPIA 183

Db 124 DEMAKYDLPATIDFILLKTKGQDKLHYVHSGTITGFIAPSTWPELAOKIKMYFALAP 183

Qy 184 TVKHAQSPGTFKLLPDMNMIKGLFGKKEFLYQTRFLRQ-LVIVLCGGVILDOICSNIMLL 242

Db 184 TVKYTETLLNKLMLVPSFLPKLIFGNKIF-YPHHFFDQFLATEVCSRETVDLLCSNALFI 242

Qy 243 LGGFNTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNSGELAFDWSGSETKNLEKCNQPT 302

Db 243 ICGFDTNNLNSRLDVTYVSHNPAGTSVQNLHWSQAVNSGELAFDWSGSETKNLEKCNQPT 302

Qy 303 TPVRYRDMVTPTAMTGGODWLSNPDVWMLSEVNTLYHKNIPFWAHVDFIWLGLDAP 362

Db 303 MPPTVNTLDMHVPITAVWSGGNDLADPDVLLPKULSNLYHKILPYNHLDPIFMANAP 362

QY 363 PHRYNEIHLHM 374
| : : : : :
Db 363 POAVNEIVSM 374
| : : : : :
RESULT 8
US-09-073-674-3
; Sequence 3, Application US/09073674
; Patent No. 5988189
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Warner-Lambert Company
; STREET: 2800 Plymouth Road
; CITY: Ann Arbor
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 48105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,674
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Crissey, Todd M.
; REGISTRATION NUMBER: 37,807
; REFERENCE/DOCKET NUMBER: 5072-D1-66-TWC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 734 622-7530
; TELEFAX: 734 622-1553
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-073-674-3

Query Match 55.0%; Score 1141.5; DB 2; Length 379;
Best Local Similarity 56.5%; Pred. No. 3.8e-114;
Matches 210; Conservative 67; Mismatches 90; Indels 5; Gaps 3;
QY 4 PTKAVDPEAFNISEIIHQGYPCPEEVEVATEDGYILSVNRIPRGLVQPKTGRPVLL 63
| : : : : :
Db 7 PT---NPEVTWNISQMTIYNGYPABEYEVVTEGYILGIDRIFYGRKNSENIGRRPVAF 63
| : : : : :
QY 64 QHGLVGASNWSINLNNLSLAFILADAGYDVLGNSRGNTWARRNLYSPDSVEFWAFSP 123
| : : : : :
Db 64 QHGLLASATNWSINLNNLSLAFILADAGYDVLGNSRGNTWARRNLYSPDSVEFWAFSP 123
| : : : : :
QY 124 DEMARFDPFVAVINFILOKTGQEKIYVGYSGTGMFTAFSTMPBLAQIKMYPALAPIA 183
| : : : : :
Db 124 DEMAYDLPATIDFILKKTGDKLHYVGHSGTGTIGTAFSTNPKLAKRIKTFYALAPVA 183
| : : : : :
QY 184 TVKHAKSPGTKEFLLPDMWIKGLFGKKEFLYQTRFLRQ-LVLYLGGQVILDOICSNIMLL 242
| : : : : :
Db 184 TVKYTETLLNKLMLVPFELFKLIFGNKIF-YPHFFDQFLATEVCSRETVDLLCSNALFI 242
| : : : : :
QY 243 LGGFTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNSGELRAPDWSGSETKNLEKCNQP 302
| : : : : :
Db 243 ICGFDTMNLNSRLDVLVLSHPAGTSVQNLHWSQAVNSGELRAPDWSGSETKNLEKCNQP 302
| : : : : :
QY 303 TPVRYRVDMTVPTAMWTGGQDMLSNPEDIYKMLSEVTNLIYHKNIPWAHVDFIWLDA 362
| : : : : :
Db 303 MPPYTNLDMHVPVIAVWNGNDLLADPHDVLKLLSKLPNLVYHKKIPPNHLDFTWAMDA 362
| : : : : :
QY 363 PHRYNEIHLHM 374
| : : : : :
Db 363 POAVNEIVSM 374
| : : : : :

RESULT 10
US-08-227-108-5
; Sequence 5, Application US/08227108
; Patent No. 5807726
; GENERAL INFORMATION:

Db 303 MPPYTNLDMHVPVIAVWNGNDLLADPHDVLKLLSKLPNLVYHKKIPPNHLDFTWAMDA 362
| : : : : :
QY 363 PHRYNEIHLHM 374
| : : : : :
Db 363 POAVNEIVSM 374
| : : : : :
RESULT 9
US-09-348-930A-2
; Sequence 2, Application US/09348930A
; Patent No. 6573431
; GENERAL INFORMATION:
; APPLICANT: Lenee, P
; APPLICANT: Gruber, V
; APPLICANT: Baudino, S
; APPLICANT: Merot, B
; APPLICANT: Benicourt, C
; APPLICANT: Cudrey, C
; TITLE OF INVENTION: Recombinant Preduodenal Lipases and Polypeptides Derivatives Produced Therefrom
; FILE REFERENCE: 18433/2012
; CURRENT APPLICATION NUMBER: US/09/348,930A
; CURRENT FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 08/945,321
; PRIOR FILING DATE: 1998-02-12
; PRIOR APPLICATION NUMBER: PCT/FR96/00606
; PRIOR FILING DATE: 1996-03-19
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 2
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Canis familiaris
US-09-348-930A-2

Query Match 55.0%; Score 1141.5; DB 4; Length 379;
Best Local Similarity 56.5%; Pred. No. 3.8e-114;
Matches 210; Conservative 67; Mismatches 90; Indels 5; Gaps 3;
QY 4 PTKAVDPEAFNISEIIHQGYPCPEEVEVATEDGYILSVNRIPRGLVQPKTGRPVLL 63
| : : : : :
Db 7 PT---NPEVTWNISQMTIYNGYPABEYEVVTEGYILGIDRIFYGRKNSENIGRRPVAF 63
| : : : : :
QY 64 QHGLVGASNWSINLNNLSLAFILADAGYDVLGNSRGNTWARRNLYSPDSVEFWAFSP 123
| : : : : :
Db 64 QHGLLASATNWSINLNNLSLAFILADAGYDVLGNSRGNTWARRNLYSPDSVEFWAFSP 123
| : : : : :
QY 124 DEMARFDPFVAVINFILOKTGQEKIYVGYSGTGMFTAFSTMPBLAQIKMYPALAPIA 183
| : : : : :
Db 124 DEMAYDLPATIDFILKKTGDKLHYVGHSGTGTIGTAFSTNPKLAKRIKTFYALAPVA 183
| : : : : :
QY 184 TVKHAKSPGTKEFLLPDMWIKGLFGKKEFLYQTRFLRQ-LVLYLGGQVILDOICSNIMLL 242
| : : : : :
Db 184 TVKYTETLLNKLMLVPFELFKLIFGNKIF-YPHFFDQFLATEVCSRETVDLLCSNALFI 242
| : : : : :
QY 243 LGGFTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNSGELRAPDWSGSETKNLEKCNQP 302
| : : : : :
Db 243 ICGFDTMNLNSRLDVLVLSHPAGTSVQNLHWSQAVNSGELRAPDWSGSETKNLEKCNQP 302
| : : : : :
QY 303 TPVRYRVDMTVPTAMWTGGQDMLSNPEDIYKMLSEVTNLIYHKNIPWAHVDFIWLDA 362
| : : : : :
Db 303 MPPYTNLDMHVPVIAVWNGNDLLADPHDVLKLLSKLPNLVYHKKIPPNHLDFTWAMDA 362
| : : : : :
QY 363 PHRYNEIHLHM 374
| : : : : :
Db 363 POAVNEIVSM 374
| : : : : :

APPLICANT: Blanchard, Claire
APPLICANT: Benicourt, Claude
APPLICANT: Junien, Jean-Louis
TITLE OF INVENTION: Recombinant Dog Gastric Lipase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/227,108
FILING DATE: 03-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fanucci, Allan A.
REGISTRATION NUMBER: 30,256
REFERENCE/DOCKET NUMBER: 7620-033
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-227-108-5

Query Match	55.0%;	Score	1141.5;	DB 1;	Length	380;			
Best Local Similarity	56.5%;	Fred.	No. 3.8e-114;						
Matches	210;	Conservative	67;	Mismatches	90;	Indels	5;	Gaps	3

QY	4	PTKAVDPFAFMNISEIIHQGYPCBEVEVATEDGYILSNVRIGLVQPKKTGSRPFWLL	63
DB	8	PT--NPEVTNNISOMITTYWGYPAEEVEVTEGDTILGIDRIIPYGRKNSENIGRRPAPFL	64
QY	64	QHGLVGGASNWTISNLPNNSLGFIILADAGFVMMGNSRGNASWKHTLSIDODETWARSY	123
DB	65	QHGLLASATNWTISNLPNNSLAFILADAGYDVLGNSRGNWTWARRNLYYSPDSVEFWASF	124
QY	124	DEMARFDLPAVINFILOKTGOKLYYGVYSGTTWGFIAFSTWPELAQIKMYFALAPIA	183
DB	125	DEMAKYDLPAITIDFLLKTTGDKUHYVGHSGTTIGTGFIAFTPKLAKIKTFYALAPVA	184
QY	184	TVKHAKSPOTKFLLLPDMWIKGLFGKKEFLYQTFSLRQ-LVIYLCQGVILDOITCSNIMLL	242
DB	185	TVKYTETLLANKLMLVPSFLFKLIIGNKIF-YPHFFPDQFLATEVCSRETVDLLCSNALPI	243
QY	243	LGGFNTNNMWSRASVYAAHTLAGTSVONILHWSQAVNSGELRAFPWGSSETKNLEKCNQP	302
DB	244	ICGFDTMLNMSRLDVIYLSHNPAGTSVQNVILHWSQAVKSGKFOAFPWGSFPVQNMHHYQS	303
QY	303	TPVRYRVRDWTPTTAMWTGGODWLSNPBDVVKMLSEVTNLIYHKNIPFWARVDFTWGLDA	362
DB	304	MPPEYNLTDMHVPVIAVWNGGNDLLADPHVDLLLSKLPNLIYRKPFPYNNLDFIWMADA	363
QY	363	PHRMVNEIIHLM	374
DB	364	POAVYNEIVSM	375

RESULT 11
US-09-073-674-5
; Sequence 5; Application US/09073674

RESULT 12

; Patent No. 5998189
 ;
 ; GENERAL INFORMATION:
 ; APPLICANT: Blanchard, Claire
 ; APPLICANT: Benicourt, Claude
 ; APPLICANT: Junien, Jean-Louis
 ; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Warner-Lambert Company
 ; STREET: 2800 Plymouth Road
 ; CITY: Ann Arbor
 ; STATE: Michigan
 ; COUNTRY: U.S.A.
 ; ZIP: 48105
 ;
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/073,674
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Crissey, Todd M.
 ; REGISTRATION NUMBER: 37,807
 ; REFERENCE/DOCKET NUMBER: 5072-D1-66-TWC
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 734 622-7530
 ; TELEFAX: 734 622-1553
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 360 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-073-674-5

Query Match	55.0%;	Score 1141.5;	DB 2;	Length 380;
Best Local Similarity	56.5%;	Pred. No. 3.8e-114;		
Matches 210;	Conservative 67;	Mismatches 90;	Indels 5;	Gaps 3;
Qy	4	PTKAVDPAEFANWISIIIOHGYPCEEVEVATGDGYILSVNRPGLVQPKKTGRSPVVLL	63	
Db	8	PT---NPETVWNISQMTYYGYPAEYEVVETDGYILGIDRIPYGRKNSENIGRPVAPL	64	
Qy	64	QHGLYGGASNWISNLPNNSLGFILADAGFDVVMGNSRGNAWSRGKHTLISLDODEFWAFSPY	123	
Db	65	QHGLLASATNWSNLPNNSLAFILADAGDYVVLGNSRGNTWARRLLYSPDSVEFWAFSF	124	
Qy	124	DEMARPDPAPVNTILQKTQEKIYYVGYSGQTTWGTAFSTMPELAQIKMYFALAPIA	183	
Db	125	DEMAQYDFEATYDILFKKTQODKXLHYVGHSGQTTIGTAFSTNPKLAKRIITFVALPVA	184	
Qy	184	TVYHAKSPGTFKLLPDMMTKGLFGKKEFLYQTRFLRQ-LVLYLGGVILDQIGCSNIMLL	242	
Db	185	TVKYTETLLNKLMLVPSPFLKIFGNKIP-YPHHFDQFLATEVCSRETVDLLCSNALFI	243	
Qy	243	LGGFNTNNMNSRASVYAAHTLAGTSVQNILHWSQAVNSGEURAPDWSGETKNLEKCNQP	302	
Db	244	ICGFDTNMLNMRDLVDVYLSHPAGTSVQNYLHWSQAVKSGKQPAFDWGSVPQNMWHYQS	303	
Qy	303	TPVRVVRDMVTVPAMWTGGDMLSNPDVGMLLSEVTNLIYHKNI PSWAHVDFPTWGLDA	362	
Db	304	MEPYNTLTKMVPATVAVWNGGNDLLADPHDVLILLSKLPNLIYHRKIPPNYHLDTIWMDA	363	
Qy	363	PHRMYNEIIHLM	374	
Db	364	POAVYNEIVSNW	375	

RESULT 12

US-09-186-489-2
; Sequence 2, Application US/09186489
; Patent No. 6375947
; GENERAL INFORMATION:
; APPLICANT: Bolen, Paul L.
; APPLICANT: Cihak, Paul L.
; APPLICANT: Scharpf Jr., Lewis G.
; TITLE OF INVENTION: Purified Recombinant Kid Pregastric Esterase, and
; TITLE OF INVENTION: Processes for its Production and Use
; FILE REFERENCE: 5499/3
; CURRENT APPLICATION NUMBER: US/09/186,489
; CURRENT FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Kid (Goat)
US-09-186-489-2

Query Match 54.3%; Score 1127; DB 4; Length 378;
Best Local Similarity 55.1%; Pred. No. 1.4e-112;
Matches 205; Conservative 66; Mismatches 101; Indels 0; Gaps 0;

Qy 7 AVDPFAFNISEIHOHQYPCPEEVEVATEDGYILSVNRIPRGLVQPKTGRSPVVLQHG 66
Db 6 AKNPEASNNVSQMSIFWGYPSEMKVITADGYILQVYRIPHGKNDANHLQREVFVLOHG 65
Qy 67 LVGGASNNISLNNPNSLGLFIADAGFDVWMCNSRGNASRKHKTLSIDQDEFWAFSYDEM 126
Db 66 LLASATNWSNLNPNLSGLFIADAGYDVLGNSRGNTWAQEHLYSPDSEFWAFSDEM 125
Qy 127 ARPLDPAVINFLQKGOEKIYVYVYSGQTTMGTFIAPSTMPPELAQKIMVFALAPIATVK 186
Db 126 AEYDLPSTIDFLKRTGQKXLYVGHSGQTTIGVAFSTNPTLAEKIEVFHALAPVATVK 185
Qy 187 HAKSPGTFKLLPDMMIKGLFKKEFLYQTRFLRQLVYLQGVQVILDOICSNIMLLGGF 246
Db 186 HTQSLFNKLALIPFLPKIIFGNKMFYHPNFEEQFLGVCSRETLDVLCKNALFAITGA 245
Qy 247 NTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNSGELRAFDWGSSETKNEKCNQPTPVR 306
Db 246 DNKNFNMSRLDYYVAHNPAGASVQNLHWRQAIKSGKFOAFDWGASVENLMHYNQPTPI 305
Qy 307 YRVDMTPTTAMWTGGQDLNPNEDVKMLSEVTNLIYHKNIPWAHVDFIWLGLDAPHRM 366
Db 306 YNLTA MVPIAVWSAGQDLADPDQVDLLSKLSNLHKEIPNYNHLDFINWADAPQEV 365
Qy 367 YNEIHLMQQEE 378
Db 366 YNEIISLMKDK 377

RESULT 13
US-10-043-665B-2
; Sequence 2, Application US/10043665B
; Patent No. 6582948
; GENERAL INFORMATION:
; APPLICANT: Bolen, Paul L.
; APPLICANT: Cihak, Paul L.
; APPLICANT: Scharpf Jr., Lewis G.
; TITLE OF INVENTION: Recombinant Kid Pregastric Esterase and Methods for Its
; TITLE OF INVENTION: Production and Use
; FILE REFERENCE: IFF-0009
; CURRENT APPLICATION NUMBER: US/10/043,665B
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 09/186,489
; PRIOR FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 378
; TYPE: PRT

; ORGANISM: Kid (Goat)
US-10-043-665B-2

Query Match 54.3%; Score 1127; DB 4; Length 378;
Best Local Similarity 55.1%; Pred. No. 1.4e-112;
Matches 205; Conservative 66; Mismatches 101; Indels 0; Gaps 0;

Qy 7 AVDPFAFNISEIHOHQYPCPEEVEVATEDGYILSVNRIPRGLVQPKTGRSPVVLQHG 66
Db 6 AKNPEASNNVSQMSIFWGYPSEMKVITADGYILQVYRIPHGKNDANHLQREVFVLOHG 65
Qy 67 LVGGASNNISLNNPNSLGLFIADAGFDVWMCNSRGNASRKHKTLSIDQDEFWAFSYDEM 126
Db 66 LLASATNWSNLNPNLSGLFIADAGYDVLGNSRGNTWAQEHLYSPDSEFWAFSDEM 125
Qy 127 ARFDLPVAVINFLQKGOEKIYVYVYSGQTTMGTFIAPSTMPPELAQKIMVFALAPIATVK 186
Db 126 AEYDLPSTIDFLKRTGQKXLYVGHSGQTTIGVAFSTNPTLAEKIEVFHALAPVATVK 185
Qy 187 HAKSPGTFKLLPDMMIKGLFKKEFLYQTRFLRQLVYLQGVQVILDOICSNIMLLGGF 246
Db 186 HTQSLFNKLALIPFLPKIIFGNKMFYHPNFEEQFLGVCSRETLDVLCKNALFAITGA 245
Qy 247 NTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNSGELRAFDWGSSETKNEKCNQPTPVR 306
Db 246 DNKNFNMSRLDYYVAHNPAGASVQNLHWRQAIKSGKFOAFDWGASVENLMHYNQPTPI 305
Qy 307 YRVDMTPTTAMWTGGQDLNPNEDVKMLSEVTNLIYHKNIPWAHVDFIWLGLDAPHRM 366
Db 306 YNLTA MVPIAVWSAGQDLADPDQVDLLSKLSNLHKEIPNYNHLDFINWADAPQEV 365
Qy 367 YNEIHLMQQEE 378
Db 366 YNEIISLMKDK 377

RESULT 14
US-09-348-930A-9
; Sequence 9, Application US/09348930A
; Patent No. 6573431
; GENERAL INFORMATION:
; APPLICANT: Lenee, P
; APPLICANT: Gruber, V
; APPLICANT: Baudino, S
; APPLICANT: Merot, B
; APPLICANT: Benicourt, C
; APPLICANT: Cudrey, C
; TITLE OF INVENTION: Recombinant Preduodenal Lipases and Polypeptides Derivatives Prod
; TITLE OF INVENTION: Plants, Processes for Obtaining Them and Their Uses
; FILE REFERENCE: 18433/2012
; CURRENT APPLICATION NUMBER: US/09/348,930A
; CURRENT FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 08/945,321
; PRIOR FILING DATE: 1998-02-12
; PRIOR APPLICATION NUMBER: PCT/FR96/00606
; PRIOR FILING DATE: 1996-03-19
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patent In version 3.0
; SEQ ID NO 9
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-348-930A-9

Query Match 53.4%; Score 1109; DB 4; Length 398;
Best Local Similarity 54.7%; Pred. No. 1.3e-110;
Matches 202; Conservative 69; Mismatches 98; Indels 0; Gaps 0;

Qy 10 PEAFNISEIHOHQYPCPEEVEVATEDGYILSVNRIPRGLVQPKTGRSPVVLQHG 69
Db 29 PEVTNISMITYGYPNEEVEVATEDGYILEVRIPYKKNKSGNTGQRPVVFVLOHGLLA 88
Qy 70 GASNNISLNNPNSLGLFIADAGFDVWMCNSRGNASRKHKTLSIDQDEFWAFSYDEM 129

```

Db      89 SATNWSLNNPNSLAFILADAGYVWLGNSRGNTWARRNLYYSPDSVEFWAFSDEWAKY 148
      130 DLPVAVINFLQKTGQEKIYYVGYSGQTTMGFTAFSTMPBLAQKIMYFALAPIATVGHAK 189
      149 DLPATIDFIVKTKGQKQLHYVGHSGQTTIGFTAFSTNPSLAKRIKTFYALAPVATVKYTK 208
      190 SPGTFLLLPDMNMIKGLFGKKEFLYQTRFLROLVYLCGOVILDOICSNIMLLLGGEFTN 249
      209 SLINKLRFVPOSLFKEIFGDKIFYPHNFDQFLATEVCSREMLNLLCSNALFIICGFSK 268
      250 NNMNGRASVYAAHTLAGTSVQNILHWSQAVNSGELRAFDPWGSETKNLEKCKNQPTFVRVYR 309
      269 NFNTRSLDVLNHPAGTSVQNMFWHTQAVKSGKQAYDWGSPVQNRMHYDQSOPPPYNNV 328
      310 RDMVTPTAMWTGGQDWLSPEDVXMLLSEVTNLIYHKNIPEWAHVDFIWLGLDAPHRMVNE 369
      329 TMMNVPIAVWNGGKOLLADPDQVGLLLPKLPNLIYHKEIPFYNHLDPIWAMDAPQEVYND 388
      370 ITHLMQOEE 378
      389 IVSMISEDK 397

```

RESULT 15

```

US-08-227-108-16
; Sequence 16, Application US/08227108
; Patent No. 5807726
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/227,108
; FILING DATE: 03-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fanucci, Allan A.
; REGISTRATION NUMBER: 30,256
; REFERENCE/DOCKET NUMBER: 7620-033
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-227-108-16

```

```

Query Match      52.9%; Score 1098.5; DB 1; Length 380;
Best Local Similarity 54.6%; Pred. No. 1.6e-109;
Matches 202; Conservative 69; Mismatches 98; Indels 1; Gaps 1;

```

```

QY      10 PEAFMNISEIIQHGYCEYEYVATEDGYILSVNRIPIRGVQPKTKGSRPVLLQHLGLVG 69
      || |||||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db      10 PEVTMNSQMITWGYPNEEYEVWTDGYILEVNRIPYVGKNSGNTGQRPVVFLQHLGLA 69
      70 GASNWSLNNPNSLGFILADAGYVWLGNSRGNTWARRNLYYSPDSVEFWAFSDEWAK 128
      70 SATNWSLNNPNSLAFILADAGYVWLGNSRGNTWARRNLYYSPDSVEFWAFSDEWAK 129
      129 FDLPAVINFLQKTGQEKIYYVGYSGQTTMGFTAFSTMPBLAQKIMYFALAPIATVGHAK 188
      130 YDLPATIDFIVKTKGQKQLHYVGHSGQTTIGFTAFSTNPSLAKRIKTFYALAPVATVKYTK 189
      189 KSPGTFLLLPDMNMIKGLFGKKEFLYQTRFLROLVYLCGOVILDOICSNIMLLLGGEFTN 248
      190 KSLINKLRFVPOSLFKEIFGDKIFYPHNFDQFLATEVCSREMLNLLCSNALFIICGFSK 249
      249 NNMNGRASVYAAHTLAGTSVQNILHWSQAVNSGELRAFDPWGSETKNLEKCKNQPTFVRVYR 308
      250 NFNTRSLDVLNHPAGTSVQNMFWHTQAVKSGKQAYDWGSPVQNRMHYDQSOPPPYNNV 309
      309 VRDMVTPTAMWTGGQDWLSPEDVXMLLSEVTNLIYHKNIPEWAHVDFIWLGLDAPHRMVNE 368
      310 TMMNVPIAVWNGGKOLLADPDQVGLLLPKLPNLIYHKEIPFYNHLDPIWAMDAPQEVYND 369
      369 ITHLMQOEE 378
      370 IVSMISEDK 379

```

Search completed: February 19, 2004, 14:23:48
Job time : 15.4317 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2004, 14:22:29 ; Search time 28.6162 Seconds
(without alignments)
3095.051 Million cell updates/sec

Title: US-10-042-431-47

Perfect score: 2247

Sequence: 1 MLETLRSQWIVSHRMEMWLL.....IHLMQEETNLQGRCEAVL 423

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2247	100.0	423	11	US-09-759-130B-417
2	2247	100.0	423	14	US-10-042-431-47
3	2174	96.8	409	10	US-09-835-996A-21
4	2076	92.4	390	11	US-09-759-130B-419
5	2076	92.4	390	14	US-10-042-431-49
6	1991	88.6	374	10	US-09-811-825-4
7	1991	88.6	398	10	US-09-811-825-2
8	1290	57.4	399	12	US-10-051-874-93
9	1289	57.4	399	12	US-10-051-874-94
10	1289	57.4	399	11	US-09-759-130B-411
11	1289	57.4	399	12	US-10-051-874-92
12	1289	57.4	399	14	US-10-042-431-41
13	1282	57.1	392	14	US-10-003-302-4
14	1165	51.8	221	11	US-09-759-130B-420
15	1165	51.8	221	14	US-10-042-431-50

16	1164.5	51.8	395	12	US-10-015-115-98
17	1154	51.4	398	12	US-10-015-115-96
18	1141	50.8	357	12	US-10-108-260A-3705
19	1139	50.7	399	12	US-10-221-097-35
20	1136	50.6	397	12	US-10-312-088-29
21	1127	50.2	378	12	US-10-420-564-2
22	1119	49.8	398	11	US-09-759-130B-445
23	1119	49.8	398	12	US-10-015-115-94
24	1119	49.8	398	12	US-10-015-115-95
25	1119	49.8	398	14	US-10-042-431-75
26	1109	49.4	371	12	US-10-015-115-97
27	1090	48.5	398	14	US-10-056-744B-2
28	1034	46.0	401	12	US-10-015-115-32
29	1031	45.9	395	14	US-10-003-302-2
30	1006	44.8	373	12	US-10-051-874-28
31	1003	44.6	373	12	US-10-051-874-30
32	977	43.5	365	12	US-10-312-088-28
33	977	43.5	395	12	US-10-051-874-90
34	937	41.7	349	12	US-10-051-874-32
35	901.5	40.1	390	12	US-10-015-115-30
36	843	37.5	351	12	US-10-051-874-91
37	783	34.8	144	11	US-09-759-130B-422
38	783	34.8	144	14	US-10-042-431-52
39	400	17.8	127	14	US-10-056-744B-5
40	319	14.2	567	12	US-10-032-585-7720
41	301	13.4	476	15	US-10-128-714-3075
42	301	13.4	641	15	US-10-128-714-8075
43	294	13.1	431	12	US-10-304-038-6
44	209	9.3	86	12	US-09-864-408A-8008
45	171	7.6	33	11	US-09-759-130B-418

ALIGNMENTS

RESULT 1

US-09-759-130B-417
; Sequence 417, Application US/09759130B
; Publication No. US20030022279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A
; APPLICANT: Fraser, Christopher C
; APPLICANT: Sharp, John D
; APPLICANT: Barnes, Thomas S
; APPLICANT: Kirst, Susan J
; APPLICANT: Mackay, Charles R
; APPLICANT: Myers, Paul S
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; TITLE OF INVENTION: USES.
; FILE REFERENCE: MP100-5350MIM
; CURRENT APPLICATION NUMBER: US/09/759,130B
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996

```

; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/602,871
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/420,707
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 417
; LENGTH: 423
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-759-1305-417

Query Match      100.0%; Score 2247; DB 11; Length 423;
Best Local Similarity 100.0%; Pred. No. 2.6e-229;
Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLETLRQWIVSHRMEMWLLILVAYMFORNVNSVHMPKAVDPEAFMNISEIIHQGYPC 60
DB 1 MLETLRQWIVSHRMEMWLLILVAYMFORNVNSVHMPKAVDPEAFMNISEIIHQGYPC 60
QY 61 EYEVEATEDGYILSVNRIPRGLVQPKTGSRPVLLQHLVGASNWIENLNNLSGLFIL 120
DB 61 EYEVEATEDGYILSVNRIPRGLVQPKTGSRPVLLQHLVGASNWIENLNNLSGLFIL 120
QY 121 ADAGFDVWNGSNGNWSRKHKTLSIDQDEFWAFSYDENARFDLPVAVINFILOKGTQGEKI 180
DB 121 ADAGFDVWNGSNGNWSRKHKTLSIDQDEFWAFSYDENARFDLPVAVINFILOKGTQGEKI 180
QY 181 YVGYSGQTTMGFIATSTWPELAQIKMYFALAPIATVVKHAKSPGTFKLLPDMMIKGLF 240
DB 181 YVGYSGQTTMGFIATSTWPELAQIKMYFALAPIATVVKHAKSPGTFKLLPDMMIKGLF 240
QY 241 GKKEFLYQTRFLRQLVILYLCGVILDOICSNIMLLGGFTNNMNSRASVYAAHTLAGT 300
DB 241 GKKEFLYQTRFLRQLVILYLCGVILDOICSNIMLLGGFTNNMNSRASVYAAHTLAGT 300
QY 301 SVQNILHWSQAVNSGELRAPDWSGSETKNLEKCNQPTPVRYRVRDMPVPTAMWTGGQDWLS 360
DB 301 SVQNILHWSQAVNSGELRAPDWSGSETKNLEKCNQPTPVRYRVRDMPVPTAMWTGGQDWLS 360
QY 361 NPEDVKMLSEVTNLIYHKNIPEWAVHDFTWGLDAPHRMYNEIIHLMQOEETNLSQGRCE 420
DB 361 NPEDVKMLSEVTNLIYHKNIPEWAVHDFTWGLDAPHRMYNEIIHLMQOEETNLSQGRCE 420
QY 421 AVL 423
DB 421 AVL 423

```

```

RESULT 2
US-10-042-431-47
; Sequence 47, Application US/10042431
; Publication No. US20020182675A1
; GENERAL INFORMATION:
; APPLICANT: MCCARTHY, Sean A
; APPLICANT: BARNES, Thomas M
; APPLICANT: FRASER, Christopher C
; APPLICANT: SHARP, John D
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
; PREVENTIVE, THERAPEUTIC, AND OTHER USES
; FILE REFERENCE: 10147-6U2
; CURRENT FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 47
; LENGTH: 423
; TYPE: PRT

```

```

; ORGANISM: Homo sapiens
; US-10-042-431-47

Query Match      100.0%; Score 2247; DB 14; Length 423;
Best Local Similarity 100.0%; Pred. No. 2.6e-229;
Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLETLRQWIVSHRMEMWLLILVAYMFORNVNSVHMPKAVDPEAFMNISEIIHQGYPC 60
DB 1 MLETLRQWIVSHRMEMWLLILVAYMFORNVNSVHMPKAVDPEAFMNISEIIHQGYPC 60
QY 61 EYEVEATEDGYILSVNRIPRGLVQPKTGSRPVLLQHLVGASNWIENLNNLSGLFIL 120
DB 61 EYEVEATEDGYILSVNRIPRGLVQPKTGSRPVLLQHLVGASNWIENLNNLSGLFIL 120
QY 121 ADAGFDVWNGSNGNWSRKHKTLSIDQDEFWAFSYDENARFDLPVAVINFILOKGTQGEKI 180
DB 121 ADAGFDVWNGSNGNWSRKHKTLSIDQDEFWAFSYDENARFDLPVAVINFILOKGTQGEKI 180
QY 181 YVGYSGQTTMGFIATSTWPELAQIKMYFALAPIATVVKHAKSPGTFKLLPDMMIKGLF 240
DB 181 YVGYSGQTTMGFIATSTWPELAQIKMYFALAPIATVVKHAKSPGTFKLLPDMMIKGLF 240
QY 241 GKKEFLYQTRFLRQLVILYLCGVILDOICSNIMLLGGFTNNMNSRASVYAAHTLAGT 300
DB 241 GKKEFLYQTRFLRQLVILYLCGVILDOICSNIMLLGGFTNNMNSRASVYAAHTLAGT 300
QY 301 SVQNILHWSQAVNSGELRAPDWSGSETKNLEKCNQPTPVRYRVRDMPVPTAMWTGGQDWLS 360
DB 301 SVQNILHWSQAVNSGELRAPDWSGSETKNLEKCNQPTPVRYRVRDMPVPTAMWTGGQDWLS 360
QY 361 NPEDVKMLSEVTNLIYHKNIPEWAVHDFTWGLDAPHRMYNEIIHLMQOEETNLSQGRCE 420
DB 361 NPEDVKMLSEVTNLIYHKNIPEWAVHDFTWGLDAPHRMYNEIIHLMQOEETNLSQGRCE 420
QY 421 AVL 423
DB 421 AVL 423

RESULT 3
US-09-835-996A-21
; Sequence 21, Application US/09835996A
; Patent No. US20020142953A1
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Dennis
; APPLICANT: Loeb, Debra
; APPLICANT: Montgomery, Julie
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhao, Qing
; APPLICANT: Wehrman, Tom
; APPLICANT: Drmanac, Radoje
; APPLICANT: Ren, Feiyan
; APPLICANT: Qian, Xiahong
; APPLICANT: Wang, Durrui
; TITLE OF INVENTION: MATERIALS AND METHODS RELATING TO LIPID METABOLISM
; FILE REFERENCE: 28110/35915A
; CURRENT APPLICATION NUMBER: US/09/835,996A
; CURRENT FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: US 60/197,137
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: US 09/714,936
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: US 09/667,298
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 09/598,042
; PRIOR FILING DATE: 2000-06-20

```

```
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-835-996A-21

Query Match      96.8%; Score 2174; DB 10; Length 409;
Best Local Similarity 100.0%; Pred. No. 1.4e-221; Indels 0; Gaps 0;
Matches 409; Conservative 0; Mismatches 0;

QY 15 MEMLLILVAYFORNVNSVHMPKAVDPPEAFMNISEIIHQGYPCBEYEVATEDGYILS 74
DB 1 MEMLLILVAYFORNVNSVHMPKAVDPPEAFMNISEIIHQGYPCBEYEVATEDGYILS 60
QY 75 VNIIPRGVLVQPKTGSRRPVVLLQHLGVGASNWIISNLPNNSLGFILADAGFDVWNGSRG 134
DB 61 VNIIPRGVLVQPKTGSRRPVVLLQHLGVGASNWIISNLPNNSLGFILADAGFDVWNGSRG 120
QY 135 NAWSRKHKTLSDIQDQEFWAFSYDEMAREDLPAVINFILOKTQGEKIYYVGYSGQTTMGFI 194
DB 121 NAWSRKHKTLSDIQDQEFWAFSYDEMAREDLPAVINFILOKTQGEKIYYVGYSGQTTMGFI 180
QY 195 AFSTWPELAQKIKMYFALAPATATVGHAKSPGTFKLLIPDMMIKGLFGKKEFLYQTRFLRQ 254
DB 181 AFSTWPELAQKIKMYFALAPATATVGHAKSPGTFKLLIPDMMIKGLFGKKEFLYQTRFLRQ 240
QY 255 LVYLICGGVILDDQICSNIMLLGGFNTNNMNSRASVVAHTLAGTSVQNTILHWSQAVNS 314
DB 241 LVYLICGGVILDDQICSNIMLLGGFNTNNMNSRASVVAHTLAGTSVQNTILHWSQAVNS 300
QY 315 GELRAFDSGETKNEKCNQPTPVRYRVRDVTPTAMTGGQDMLSNPDEVKMLLSEVTN 374
DB 301 GELRAFDSGETKNEKCNQPTPVRYRVRDVTPTAMTGGQDMLSNPDEVKMLLSEVTN 360
QY 375 LIYHKNIPENAHVDFIWLGDAPHRMYNEIIHLMQOEETNLSQGRCEAVL 423
DB 361 LIYHKNIPENAHVDFIWLGDAPHRMYNEIIHLMQOEETNLSQGRCEAVL 409

RESULT 4
US-09-759-130B-419
; Sequence 419, Application US/09759130B
; Publication No. US2003002279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A
; APPLICANT: Fraser, Christopher C
; APPLICANT: Sharp, John D
; APPLICANT: Barnes, Thomas S
; APPLICANT: Kirst, Susan J
; APPLICANT: Mackay, Charles R
; APPLICANT: Myers, Paul S
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; FILE REFERENCE: MP100-5350NIM
; CURRENT APPLICATION NUMBER: US/09/759,130B
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US/09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US/09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US/09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US/09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US/09/596,194
```

```
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US/09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US/09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US/09/393,996
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US/09/602,871
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US/09/420,707
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 419
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-759-130B-419

Query Match      92.4%; Score 2076; DB 11; Length 390;
Best Local Similarity 100.0%; Pred. No. 3.2e-211; Indels 0; Gaps 0;
Matches 390; Conservative 0; Mismatches 0;

QY 34 VHMPTKAVDPPEAFMNISEIIHQGYPCBEYEVATEDGYILSVNRIIPRGVLVQPKTGSRRPV 93
DB 1 VHMPTKAVDPPEAFMNISEIIHQGYPCBEYEVATEDGYILSVNRIIPRGVLVQPKTGSRRPV 60
QY 94 VLLQHLGVGASNWIISNLPNNSLGFILADAGFDVWNGSRGNASRKHKTLSDIQDQEFWA 153
DB 61 VLLQHLGVGASNWIISNLPNNSLGFILADAGFDVWNGSRGNASRKHKTLSDIQDQEFWA 120
QY 154 PSYDEWARFDLPAVINFILOKTQGEKIYYVGYSGQTTMGFIASFSTWPELAQKIKMYFALA 213
DB 121 PSYDEWARFDLPAVINFILOKTQGEKIYYVGYSGQTTMGFIASFSTWPELAQKIKMYFALA 180
QY 214 PIATVGHAKSPGTFKLLIPDMMIKGLFGKKEFLYQTRFLRQLVYLICGGVILDDQICSNIM 273
DB 181 PIATVGHAKSPGTFKLLIPDMMIKGLFGKKEFLYQTRFLRQLVYLICGGVILDDQICSNIM 240
QY 274 LLLGGFNTNNMNSRASVVAHTLAGTSVQNTILHWSQAVNSGELRAFDSGETKNEKCN 333
DB 241 LLLGGFNTNNMNSRASVVAHTLAGTSVQNTILHWSQAVNSGELRAFDSGETKNEKCN 300
QY 334 QPTPVRYRVRDVTPTAMTGGQDMLSNPDEVKMLLSEVTNLIYHKNIPENAHVDFIWL 393
DB 301 QPTPVRYRVRDVTPTAMTGGQDMLSNPDEVKMLLSEVTNLIYHKNIPENAHVDFIWL 360
QY 394 DAPHRMYNEIIHLMQOEETNLSQGRCEAVL 423
DB 361 DAPHRMYNEIIHLMQOEETNLSQGRCEAVL 390

RESULT 5
US-10-042-431-49
; Sequence 49, Application US/10042431
; Publication No. US20020182675A1
; GENERAL INFORMATION:
; APPLICANT: MCCARTHY, Sean A
; APPLICANT: BARNES, Thomas M
; APPLICANT: FRASER, Christopher C
; APPLICANT: SHARP, John D
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
; TITLE OF INVENTION: PREVENTIVE, THERAPEUTIC, AND OTHER USES
; FILE REFERENCE: 10147-6U2
; CURRENT APPLICATION NUMBER: US/10/042,431
; CURRENT FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US/09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US/09/578,063
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
```

```
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-431-49

Query Match      92.4%; Score 2076; DB 14; Length 390;
Best Local Similarity 100.0%; Pred. No. 3.2e-211;
Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 VHMPTKAVDPEAFMNISEIIHQGYPCBEYEVATEDGYILSVNRI PRGLVQPKTGSRPV 93
DB 1 VHMPTKAVDPEAFMNISEIIHQGYPCBEYEVATEDGYILSVNRI PRGLVQPKTGSRPV 60

QY 94 VLLQHLGVLGASWISNLPNNSLGFILADAGFDVWNGSRGNWSRKHKTLSIDQDEFWA 153
DB 61 VLLQHLGVLGASWISNLPNNSLGFILADAGFDVWNGSRGNWSRKHKTLSIDQDEFWA 120

QY 154 PSYDEMARFDLPVINFILQKTQOEKIYYGVYSGQTTMGFIAPSTMPPELAQIKMYFALA 213
DB 121 PSYDEMARFDLPVINFILQKTQOEKIYYGVYSGQTTMGFIAPSTMPPELAQIKMYFALA 180

QY 214 PIATVXHAKEPGTKFLLPDMIKGLFGKKEFLYQTRFLRQLVYLCOGVILDOICSNIM 273
DB 181 PIATVXHAKEPGTKFLLPDMIKGLFGKKEFLYQTRFLRQLVYLCOGVILDOICSNIM 240

QY 274 LLLGGFTNNMNSRASVVAHTLAGTSVQNILHWSQAVNSGELRAFDWGSSEKCN 333
DB 241 LLLGGFTNNMNSRASVVAHTLAGTSVQNILHWSQAVNSGELRAFDWGSSEKCN 300

QY 334 QPTPVRYRVRDVTPTAMTGGQDLSNPEDVKMLSEVTNLIYHKNIPWAHVDPIWGL 393
DB 301 QPTPVRYRVRDVTPTAMTGGQDLSNPEDVKMLSEVTNLIYHKNIPWAHVDPIWGL 360

QY 394 DAPHRMYNEIHLMOEETNLSQRCCEAVL 423
DB 361 DAPHRMYNEIHLMOEETNLSQRCCEAVL 390

RESULT 6
US-09-811-825-4
; Sequence 4, Application US/09811825
; Patent No. US20020144297A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001170
; CURRENT APPLICATION NUMBER: US/09/811,825
; CURRENT FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Human
US-09-811-825-4

Query Match      88.6%; Score 1991; DB 10; Length 374;
Best Local Similarity 100.0%; Pred. No. 3.1e-202;
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 SEIIHQGYPCBEYEVATEDGYILSVNRI PRGLVQPKTGSRPVLLQHLGVLGASNWIS 109
DB 1 SEIIHQGYPCBEYEVATEDGYILSVNRI PRGLVQPKTGSRPVLLQHLGVLGASNWIS 60

QY 110 NLPNNSLGFILADAGFDVWNGSRGNWSRKHKTLSIDQDEFWAFSYDEMARFDLPV 169
DB 61 NLPNNSLGFILADAGFDVWNGSRGNWSRKHKTLSIDQDEFWAFSYDEMARFDLPV 120

QY 170 FILOKTQOEKIYYGVYSGQTTMGFIAPSTMPPELAQIKMYFALAPIATVXHAKEPGTK 229
DB 121 FILOKTQOEKIYYGVYSGQTTMGFIAPSTMPPELAQIKMYFALAPIATVXHAKEPGTK 180

RESULT 7
US-09-811-825-2
; Sequence 2, Application US/09811825
; Patent No. US20020144297A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001170
; CURRENT APPLICATION NUMBER: US/09/811,825
; CURRENT FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Human
US-09-811-825-2

Query Match      88.6%; Score 1991; DB 10; Length 398;
Best Local Similarity 100.0%; Pred. No. 3.4e-202;
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 SEIIHQGYPCBEYEVATEDGYILSVNRI PRGLVQPKTGSRPVLLQHLGVLGASNWIS 109
DB 25 SEIIHQGYPCBEYEVATEDGYILSVNRI PRGLVQPKTGSRPVLLQHLGVLGASNWIS 84

QY 110 NLPNNSLGFILADAGFDVWNGSRGNWSRKHKTLSIDQDEFWAFSYDEMARFDLPV 169
DB 85 NLPNNSLGFILADAGFDVWNGSRGNWSRKHKTLSIDQDEFWAFSYDEMARFDLPV 144

QY 170 FILOKTQOEKIYYGVYSGQTTMGFIAPSTMPPELAQIKMYFALAPIATVXHAKEPGTK 229
DB 145 FILOKTQOEKIYYGVYSGQTTMGFIAPSTMPPELAQIKMYFALAPIATVXHAKEPGTK 204

QY 230 LLLPDMIKGLFGKKEFLYQTRFLRQLVYLCOGVILDOICSNIMLLGGFTNNMNSRA 289
DB 205 LLLPDMIKGLFGKKEFLYQTRFLRQLVYLCOGVILDOICSNIMLLGGFTNNMNSRA 264

QY 290 SVYAAHTLAGTSVQNILHWSQAVNSGELRAFDWGSSEKCNQPTPVRYRVRDVTPT 349
DB 265 SVYAAHTLAGTSVQNILHWSQAVNSGELRAFDWGSSEKCNQPTPVRYRVRDVTPT 324

QY 350 AMWTGGQDLSNPEDVKMLSEVTNLIYHKNIPWAHVDPIWGLDAPHRMYNEIHLMOQ 409
DB 325 AMWTGGQDLSNPEDVKMLSEVTNLIYHKNIPWAHVDPIWGLDAPHRMYNEIHLMOQ 384

QY 410 EETNLSQRCCEAVL 423
DB 385 EETNLSQRCCEAVL 398

RESULT 8
US-10-051-874-93
```

```

; Sequence 93, Application US/10051874
; Publication No. US20040005557A1
; GENERAL INFORMATION:
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Alsobrook II, John P
; APPLICANT: Colman, Steven D
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Boldog, Ferenc
; APPLICANT: Vernet, Corine AM
; APPLICANT: Li, Li
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Casman, Stacie J
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Edinger, Shlomit R
; APPLICANT: MacDougall, John R
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Patturajan, Meera
; APPLICANT: Shinkets, Richard A
; APPLICANT: Pena, Carol EA
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Millet, Isabelle
; APPLICANT: Miller, Charles E
; APPLICANT: Lepley, Denise M
; APPLICANT: Smithson, Glenda
; APPLICANT: Baumgartner, Jason C
; APPLICANT: Herrman, John L
; APPLICANT: Peyman, John A
; APPLICANT: Gorman, Linda
; APPLICANT: Mezes, Peter D
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Gerlach, Valerie
; APPLICANT: Grosse, William M
; APPLICANT: Liu, Xiaohong
; APPLICANT: Ellerman, Karen
; APPLICANT: Rothenberg, Mark
; APPLICANT: Stone, David J
; APPLICANT: Burgess, Catherine E
; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
; TITLE OF INVENTION: USING THE SAME
; FILE REFERENCE: 21402-245
; CURRENT APPLICATION NUMBER: US/10/051,874
; CURRENT FILING DATE: 2002-09-25
; PRIOR APPLICATION NUMBER: 60/268,595
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/325,306
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/262,587
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/272,409
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/262,454
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/276,777
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/291,672
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/330,336
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/265,530
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/261,376
; PRIOR FILING DATE: 2001-01-16
; NUMBER OF SEQ ID NOS: 269
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 93
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-051-874-93

```

Query Match 57.4%; Score 1290; DB 12; Length 399;

```

Best Local Similarity 61.0%; Pred. No. 8.1e-128;
Matches 242; Conservative 59; Mismatches 94; Indels 2; Gaps 1;
Qy 15 MEMMLILVAYMPQRNVNVMPTK--AVDPEAFNMISIIHQGYPCPEEYEVATEDGYI 72
Db 1 MKORFLGLVVLVLTWLTSEGGKLTAVDPETNNMNVSEIIISYWGFPSEYLVETEDGYI 60
Qy 73 LSVNRIPLGLVOPKKTGSRPVVLQHLQGLVGGASNNISLNNNSLGFILADAGFDVMMGNS 132
Db 61 LCLNRIPLGRKSHSDKGPVVFLOHGLLADSSNVTNLANSLSGLFILADAGFDVMMGNS 120
Qy 133 RGNWRSKHKTLISIDODEFWAFSYDEMAFPDLPVAVNFILQTKGQEKIYVYGVSOQTWNG 192
Db 121 RGNWRSKHKTLISVSDQDEFWAFSYDEMAKYDLPASINFILNKTGQEQVYVGVSHSOQTWNG 180
Qy 193 FIAPSTMPPELAOKIKWYFALAPIATVTKHAKSPGTFLLLPDMWIKGLFGKKEFLYQTRFL 252
Db 181 FIAPSOIPELAKRIKMFALGPVAVFCTSPAKLGRPLDHLIKDLFGDKKEFLPQSAFL 240
Qy 253 RQVLYLCQGVILDDQICSNIMLLGQFNNTNNMNGRASVYAAHTLAGTSVQNLHWSQAV 312
Db 241 KWLGTHTVCTHVLKELCCNLCLLCCFNERNLNSRDVYVTHSPAGTSVQNLHWSQAV 300
Qy 313 NSCELAPDWSGETKNLEKCNQOPTPVVRVVRDVTMTGQDQWLSPNPDVXMLSEV 372
Db 301 KFOKQAFDWSGSAKNYFHNOSYPTTYNVKMDLVTAVSGGHDWLDADVVDNILLTQI 360
Qy 373 TNLVYHKNIPWAHVDFIWLGDAPHMYNEIHLMOQ 409
Db 361 TNLVFTHEPIWEHLDPIWGLDAPWRLYKNIINLRK 397

```

RESULT 9

```

US-10-051-874-94
; Sequence 94, Application US/10051874
; Publication No. US20040005557A1
; GENERAL INFORMATION:
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Alsobrook II, John P
; APPLICANT: Colman, Steven D
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Boldog, Ferenc
; APPLICANT: Vernet, Corine AM
; APPLICANT: Li, Li
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Casman, Stacie J
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Edinger, Shlomit R
; APPLICANT: MacDougall, John R
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Patturajan, Meera
; APPLICANT: Shinkets, Richard A
; APPLICANT: Pena, Carol EA
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Millet, Isabelle
; APPLICANT: Miller, Charles E
; APPLICANT: Lepley, Denise M
; APPLICANT: Smithson, Glenda
; APPLICANT: Baumgartner, Jason C
; APPLICANT: Herrman, John L
; APPLICANT: Peyman, John A
; APPLICANT: Gorman, Linda
; APPLICANT: Mezes, Peter D
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Gerlach, Valerie
; APPLICANT: Grosse, William M
; APPLICANT: Liu, Xiaohong
; APPLICANT: Ellerman, Karen
; APPLICANT: Rothenberg, Mark
; APPLICANT: Stone, David J
; APPLICANT: Burgess, Catherine E

```

;; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
;; TITLE OF INVENTION: USING THE SAME

;; FILE REFERENCE: 21402-245
;; CURRENT APPLICATION NUMBER: US/10/051,874
;; PRIOR FILING DATE: 2002-09-25
;; PRIOR APPLICATION NUMBER: 60/268,595
;; PRIOR FILING DATE: 2001-02-14
;; PRIOR APPLICATION NUMBER: 60/325,306
;; PRIOR FILING DATE: 2001-09-27
;; PRIOR APPLICATION NUMBER: 60/262,587
;; PRIOR FILING DATE: 2001-01-18
;; PRIOR APPLICATION NUMBER: 60/272,409
;; PRIOR FILING DATE: 2001-02-28
;; PRIOR APPLICATION NUMBER: 60/262,454
;; PRIOR FILING DATE: 2001-01-18
;; PRIOR APPLICATION NUMBER: 60/276,777
;; PRIOR FILING DATE: 2001-03-16
;; PRIOR APPLICATION NUMBER: 60/291,672
;; PRIOR FILING DATE: 2001-05-17
;; PRIOR APPLICATION NUMBER: 60/330,336
;; PRIOR FILING DATE: 2001-10-18
;; PRIOR APPLICATION NUMBER: 60/265,530
;; PRIOR FILING DATE: 2001-01-31
;; PRIOR APPLICATION NUMBER: 60/261,376
;; PRIOR FILING DATE: 2001-01-16
;; NUMBER OF SEQ ID NOS: 269
;; SOFTWARE: Patent in Ver. 2.1
;; SEQ ID NO 94
;; LENGTH: 399
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-051-874-94

Query Match 57.4%; Score 1290; DB 12; Length 399;
Best Local Similarity 61.0%; Pred. No. 8.1e-128;
Matches 242; Conservative 59; Mismatches 94; Indels 2; Gaps 1;

QY 15 MEMLLIIVAFQFNQVNSVHPTK--AVDPEAFMNISEIIHQGYPCREVEVATEDGYI 72
DB 1 MKRFLGLVCLVLTWTHSESRKLTAVDETMNNSVSEIISYWGFPSEYLVETEDGYI 60
QY 73 LSVNRIPRGLVOPKTKSRPVLVQLHGLVGASNIWNLNPNNSLGLFADAGFDVWNGNS 132
DB 61 LCLNRIPIGRKXKHSKDKGPKPVVFLQHLGLADSSNWNVTNLANSLSGLFADAGFDVWNGNS 120
QY 133 RGNWSRKHKTLSDIDQEFWAFSDEWARPDLPAVINFLIKTGOEKIYVYVGSQGTMG 192
DB 121 RGNWSRKHKTLSDIDQEFWAFSDEWAKYDLPAVINFLIKTGOEKIYVYVGSQGTMG 180
QY 193 FIATSTPELAQKIMYFALAPIATVYKHAKEPGTKFLLPDMNMIKGLFGKBEFYQTRPL 252
DB 181 FIATSTPELAQKIMYFALGFAVAPCTSPMAKLGRLPDHLIKDLFGDKFLPQSAFL 240
QY 253 RQVLYLCQVTLQICSNIMLLGGFTNMMNSRASVYAAHTLAGTSVQNLHWSQAV 312
DB 241 KWLGHVTHVILKELCNLCFLCGFERNLNSRVDVYTHSPAGTSVQNLHWSQAV 300
QY 313 NSGELRAFDFWSETKNLEKCNQPTPVRVYRVDMTVPTAMTGGQDLNPNEDVKNLSEV 372
DB 301 KFQKFAQFDWSSAKNYHYNQSYPTVNVKMLVPTAVMSGGHDVLADVYDVNILLTQI 360
QY 373 TNLVYHKNIPWAHVDFTWGLDAPHMTNVEIHLMOQ 409
DB 361 TNLVPHESIPEWEHLDFIWLGLDAPWRLYKNIINLMRK 397

RESULT 10
US-09-759-130B-411
; Sequence 411, Application US/09759130B
; Publication No. US2003002279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A

;; APPLICANT: Fraser, Christopher C
;; APPLICANT: Sharp, John D
;; APPLICANT: Barnes, Thomas S
;; APPLICANT: Kirt, Susan J
;; APPLICANT: Mackay, Charles R
;; APPLICANT: Myers, Paul S
;; APPLICANT: Leiby, Kevin R
;; APPLICANT: Wrighton, Nicolas
;; APPLICANT: Goodearl, Andrew
;; APPLICANT: Holtzman, Douglas A
;; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
;; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
;; FILE REFERENCE: MPI00-5350NMIM
;; CURRENT APPLICATION NUMBER: US/09/759,130B
;; CURRENT FILING DATE: 2002-09-16
;; PRIOR APPLICATION NUMBER: US 09/479,249
;; PRIOR FILING DATE: 2000-01-07
;; PRIOR APPLICATION NUMBER: US 09/559,497
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: US 09/578,063
;; PRIOR FILING DATE: 2000-05-24
;; PRIOR APPLICATION NUMBER: US 09/333,159
;; PRIOR FILING DATE: 1999-06-14
;; PRIOR APPLICATION NUMBER: US 09/596,194
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: US 09/342,364
;; PRIOR FILING DATE: 1999-06-29
;; PRIOR APPLICATION NUMBER: US 09/608,452
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/393,996
;; PRIOR FILING DATE: 1999-09-10
;; PRIOR APPLICATION NUMBER: US 09/602,871
;; PRIOR FILING DATE: 2000-06-23
;; PRIOR APPLICATION NUMBER: US 09/420,707
;; PRIOR FILING DATE: 1999-10-19
;; NUMBER OF SEQ ID NOS: 460
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 411
;; LENGTH: 399
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-759-130B-411

Query Match 57.4%; Score 1289; DB 11; Length 399;
Best Local Similarity 63.5%; Pred. No. 1e-127;
Matches 235; Conservative 55; Mismatches 80; Indels 0; Gaps 0;

QY 40 AVDPAPFNISEIIHQGYPCREVEVATEDGYILSVNRIPRGLVOPKTKSRPVLVQLHG 99
DB 28 AVDPETNMNVEIISYWGFPSEYLVETEDGYILCLNRIPIGRKXKHSKDKGPKPVVFLQHG 87
QY 100 LVGGASNIWNLNPNNSLGLFADAGFDVWNGNSRGNWSRKHKTLSDIDQEFWAFSDE 159
DB 88 LQADSSNWNVTNLANSLSGLFADAGFDVWNGNSRGNWSRKHKTLSDIDQEFWAFSDE 147
QY 160 AFEDLPVAVINFLIKTGOEKIYVYVGSQGTMGFIATSTPELAQKIMYFALAPIATVK 219
DB 148 AKYDLPASINFLIKTGOEKIYVYVGSQGTMGFIATSTPELAQKIMYFALAPIATVK 207
QY 220 HAKSPGTFLLLPDMNMIKGLFGKBEFYQTRFLQVLYLCQVTLQICSNIMLLGGF 279
DB 208 FCTSPMAKLGRLPDHLIKDLFGDKFLPQSAFLKWLGHVTHVILKELCNLCFLCGF 267
QY 280 NTNNNNMNSRASVYAAHTLAGTSVQNLHWSQAVNLSGELRAFDFWSETKNLEKCNQPTPVR 339
DB 268 NERNLNSRVDVYTHSPAGTSVQNLHWSQAVKFAQFDWSSAKNYHYNQSYPT 327
QY 340 YAVRDMTVPTAMTGGQDLNPNEDVKNLSEVTLNLYHKNIPWAHVDFTWGLDAPH 399
DB 328 YNVKMLVPTAVMSGGHDVLADVYDVNILLTQITNLVPHESIPEWEHLDFIWLGLDAPWRL 387
QY 400 YNEIHLMOQ 409

Db 388 YNKIINLMRK 397

RESULT 11

```

RESULT 11
US-10-051-874-92
/ Sequence 92, Application US/10051874
/ Publication No. US20040005557A1
/ GENERAL INFORMATION:
/ APPLICANT: Padigar, Muralidhara
/ APPLICANT: Alsebrook II, John P
/ APPLICANT: Colman, Steven D
/ APPLICANT: Spytek, Kimberly A
/ APPLICANT: Boldog, Ferenc
/ APPLICANT: Vernet, Corine AM
/ APPLICANT: Li, Li
/ APPLICANT: Shenoy, Suresh G
/ APPLICANT: Casman, Stacie J
/ APPLICANT: Guo, Xiaojia Sasha
/ APPLICANT: Edinger, Shomit R
/ APPLICANT: Macdougall, John R
/ APPLICANT: Mallyankar, Uriel M
/ APPLICANT: Patturajan, Meera
/ APPLICANT: Shinkets, Richard A
/ APPLICANT: Pena, Carol EA
/ APPLICANT: Tchernev, Velizar T
/ APPLICANT: Zerkusen, Bryan D
/ APPLICANT: Millet, Isabelle
/ APPLICANT: Miller, Charles E
/ APPLICANT: Lepley, Denise M
/ APPLICANT: Smithson, Glenda
/ APPLICANT: Baumgartner, Jason C
/ APPLICANT: Herrman, John L
/ APPLICANT: Peyman, John A
/ APPLICANT: Gorman, Linda
/ APPLICANT: Mezes, Peter D
/ APPLICANT: Kekuda, Ramesh
/ APPLICANT: Taulier Jr, Raymond J
/ APPLICANT: Gerlach, Valerie
/ APPLICANT: Grosse, William M
/ APPLICANT: Liu, Xiaohong
/ APPLICANT: Ellerman, Karen
/ APPLICANT: Rothenberg, Mark
/ APPLICANT: Stone, David J
/ APPLICANT: Burgess, Catherine E
/ TITLE OF INVENTION: PROTEINS, POLYNT
/ TITLE OF INVENTION: USING THE SAME
/ FILE REFERENCE: 21402-245
/ CURRENT APPLICATION NUMBER: US/10/0
/ CURRENT FILING DATE: 2002-09-25
/ PRIOR APPLICATION NUMBER: 60/268,591
/ PRIOR FILING DATE: 2001-02-14
/ PRIOR APPLICATION NUMBER: 60/325,301
/ PRIOR FILING DATE: 2001-09-27
/ PRIOR APPLICATION NUMBER: 60/262,581
/ PRIOR FILING DATE: 2001-01-18
/ PRIOR APPLICATION NUMBER: 60/272,401
/ PRIOR FILING DATE: 2001-02-28
/ PRIOR APPLICATION NUMBER: 60/262,451
/ PRIOR FILING DATE: 2001-01-18
/ PRIOR APPLICATION NUMBER: 60/276,771
/ PRIOR FILING DATE: 2001-03-16
/ PRIOR APPLICATION NUMBER: 60/291,671
/ PRIOR FILING DATE: 2001-05-17
/ PRIOR APPLICATION NUMBER: 60/330,331
/ PRIOR FILING DATE: 2001-10-18
/ PRIOR APPLICATION NUMBER: 60/265,531
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/261,371
/ PRIOR FILING DATE: 2001-01-16
/ NUMBER OF SEQ ID NOS: 269
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 92

```

```

Db 88 LADSNWNTNLSNLSGLADAGFDVWNGSRGNTSRKHTLSVSDQDFWAFSYDEM 147
QY 160 ARPDLPVAVINFILOKQTKQYIYVGSQGTGTFIAPSTMPBLAQKIRYFALAPIATVK 219
Db 148 AKYDLFASINFLNKTGQYVYVGHSGQTTGTFIAPSQIPBLAKRIRKMFALGFVASVA 207
QY 220 HAKSPGTEKLLPDMMIKGLGKKEFLYQTRFLRQVLYLQGVILDOICSNIMLLGCF 279
Db 208 FCTSPWAKLGRUPDLHILKDLGKFLPOSFLKWLGHVCHVILKELCNLCFLGCF 267
QY 280 NTNNMMSRASYAAHTLAGTSVQNLHWSQAVNSGELRAFDWGSSEKTKNLEKCNQPTVR 339
Db 268 NERNLNMSRVDVYTHSPAGTSVQNNLHWSQAVKQKQAFDWGSSAKNYFYHNGSYPT 327
QY 340 YVRDVTPTAMTGGQDMLSPEDVKMLLSVTLNLYHKNIPWAHVDFFIWLGDAPHRM 399
Db 328 YNVKMLVPTAVWSGGHDLADVDVNNILLTQITNLVFHESIPWEHLDFFIWLGDAPWRL 387
QY 400 YNEIHLMOQ 409
Db 388 YNKIINLMK 397

```

```

RESULT 13
US-10-003-302-4
; Sequence 4, Application US/10003302
; Publication No. US20020142435A1
; GENERAL INFORMATION:
; APPLICANT: MERKULOV, Gennady et al.
; TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES
; TITLE OF INVENTION: THERIOF
; FILE REFERENCE: CL001186DIV
; CURRENT APPLICATION NUMBER: US/10/003,302
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-003-302-4

```

```

Query Match 57.1%; Score 1282; DB 14; Length 392;
Best Local Similarity 63.8%; Pred. No. 5.5e-127; Indels 0; Gaps 0;
Matches 234; Conservative 53; Mismatches 80;

QY 40 AVDPETNNVSEIISYWGPPSEYLVETEDGYILCLNRIPIHGRKNHSDGPKPVVFLQH 85
Db 26 AVDPETNNVSEIISYWGPPSEYLVETEDGYILCLNRIPIHGRKNHSDGPKPVVFLQH 85
QY 100 LVGASGNWISNLPNSLGFILADAGFDVWNGSRGNMRSRKHKTLSDQDFWAFSYDEM 159
Db 86 LIADSNWNTNLSNLSGLADAGFDVWNGSRGNTSRKHTLSVSDQDFWAFSYDEM 145
QY 160 ARPDLPVAVINFILOKQTKQYIYVGSQGTGTFIAPSTMPBLAQKIRYFALAPIATVK 219
Db 146 AKYDLFASINFLNKTGQYVYVGHSGQTTGTFIAPSQIPBLAKRIRKMFALGFVASVA 205
QY 220 HAKSPGTEKLLPDMMIKGLGKKEFLYQTRFLRQVLYLQGVILDOICSNIMLLGCF 279
Db 206 FCTSPWAKLGRUPDLHILKDLGKFLPOSFLKWLGHVCHVILKELCNLCFLGCF 265
QY 280 NTNNMMSRASYAAHTLAGTSVQNLHWSQAVNSGELRAFDWGSSEKTKNLEKCNQPTVR 339
Db 266 NERNLNMSRVDVYTHSPAGTSVQNNLHWSQAVKQKQAFDWGSSAKNYFYHNGSYPT 325
QY 340 YVRDVTPTAMTGGQDMLSPEDVKMLLSVTLNLYHKNIPWAHVDFFIWLGDAPHRM 399
Db 326 YNVKMLVPTAVWSGGHDLADVDVNNILLTQITNLVFHESIPWEHLDFFIWLGDAPWRL 385
QY 400 YNEIHL 406

```

```

Db 386 YNKIINL 392
RESULT 14
US-09-759-130B-420
; Sequence 420, Application US/09759130B
; Publication No. US20030022279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A
; APPLICANT: Fraser, Christopher C
; APPLICANT: Sharp, John D
; APPLICANT: Barnes, Thomas S
; APPLICANT: Kirt, Susan J
; APPLICANT: Mackay, Charles R
; APPLICANT: Myers, Paul S
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; TITLE OF INVENTION: USES
; FILE REFERENCE: MP100-5350NMIM
; CURRENT APPLICATION NUMBER: US/09/759,130B
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/602,871
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/420,707
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 420
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-759-130B-420

```

```

Query Match 51.8%; Score 1165; DB 11; Length 221;
Best Local Similarity 100.0%; Pred. No. 6e-115;
Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 VHMPTKAVDPAPMNISEIIHQGYPCEEYEVATEDGYILSVNRIPIRGLVQPKTGSREV 93
Db 1 VHMPTKAVDPAPMNISEIIHQGYPCEEYEVATEDGYILSVNRIPIRGLVQPKTGSREV 60
QY 94 VLLQHLGVGASNWIENLNNLSLGFILADAGFDVWNGSRGNMRSRKHKTLSDQDFWA 153
Db 61 VLLQHLGVGASNWIENLNNLSLGFILADAGFDVWNGSRGNMRSRKHKTLSDQDFWA 120
QY 154 FSDVEMARFDLPVAVINFILOKQTKQYIYVGSQGTGTFIAPSTMPBLAQKIRYFAL 213
Db 121 FSDVEMARFDLPVAVINFILOKQTKQYIYVGSQGTGTFIAPSTMPBLAQKIRYFAL 180
QY 214 PIATVHKASPGTKFLLLPDMMIKGLGKKEFLYQTRFLRQ 254

```


Db 181 PIATVHAKSPGTFKLLPDMNKGKLGKKEFLYOTRFLRQ 221

RESULT 15
 US-10-042-431-50
 ; Sequence 50, Application US/10042431
 ; Publication No. US20020182675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MCCARTHY, Sean A
 ; APPLICANT: BARNES, Thomas M
 ; APPLICANT: FRASER, Christopher C
 ; APPLICANT: SHARP, John D
 ; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
 ; FILE REFERENCE: 10147-6U2
 ; CURRENT APPLICATION NUMBER: US/10/042,431
 ; CURRENT FILING DATE: 2001-10-25
 ; PRIOR APPLICATION NUMBER: US 09/333,159
 ; PRIOR FILING DATE: 1999-06-14
 ; PRIOR APPLICATION NUMBER: US 09/578,063
 ; PRIOR FILING DATE: 2000-05-24
 ; NUMBER OF SEQ ID NOS: 79
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 50
 ; LENGTH: 221
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-042-431-50

Query Match 51.8%; Score 1165; DB 14; Length 221;
 Best Local Similarity 100.0%; Pred. No. 6e-115;
 Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 34 VHMPTKAVDPEAFMNISEIIHQGYPCBEYEVATEDGYILSVNRIPRGLVQPKTGRPV 93
 DB 1 VHMPTKAVDPEAFMNISEIIHQGYPCBEYEVATEDGYILSVNRIPRGLVQPKTGRPV 60
 QY 94 VLLQHLVCGAGNWSNLNPNLSGLFILADAGPDVWNGSRGNWGRKHTLSIDQDEFWA 153
 DB 61 VLLQHLVCGAGNWSNLNPNLSGLFILADAGPDVWNGSRGNWGRKHTLSIDQDEFWA 120
 QY 154 FSYDEMARFDLPAVINFILOKTQGEKIYVYVGSQGTWGFIAFSTMPFLAQKIMYPALA 213
 DB 121 FSYDEMARFDLPAVINFILOKTQGEKIYVYVGSQGTWGFIAFSTMPFLAQKIMYPALA 180
 QY 214 PIATVHAKSPGTFKLLPDMNKGKLGKKEFLYOTRFLRQ 254
 DB 181 PIATVHAKSPGTFKLLPDMNKGKLGKKEFLYOTRFLRQ 221

Search completed: February 19, 2004, 14:34:24
 Job time : 30.6162 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2004, 14:19:24 ; Search time 14.5683 Seconds
(without alignments)
1228.526 Million cell updates/sec

Title: US-10-042-431-47
Perfect score: 2247
Sequence: 1 MLTSLRQWIVSHRMENMLL.....IHLMQQETNLSQRCCEAVL 423

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCFUS.COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1282	57.1	392	4	US-09-820-001-4
2	1161	51.7	377	1	US-08-227-108-17
3	1161	51.7	377	1	US-09-073-674-17
4	1150.5	51.2	379	1	US-08-227-108-18
5	1150.5	51.2	379	2	US-09-073-674-18
6	1141.5	50.8	375	4	US-09-348-930A-6
7	1141.5	50.8	379	1	US-08-227-108-3
8	1141.5	50.8	379	4	US-09-348-930A-2
9	1141.5	50.8	380	1	US-08-227-108-5
10	1141.5	50.8	380	2	US-09-073-674-5
11	1141.5	50.8	380	1	US-09-186-489-2
12	1127	50.2	378	4	US-10-043-665B-2
13	1119	49.8	398	4	US-09-348-930A-9
14	1098.5	48.9	380	1	US-08-227-108-16
15	1098.5	48.9	380	2	US-09-073-674-16
16	1098.5	48.9	380	1	US-09-820-001-2
17	1031	45.9	395	4	US-09-348-930A-4
18	1003	44.6	325	4	US-09-252-991A-30091
19	143.5	6.4	348	4	US-08-484-105-22
20	105	4.7	430	1	US-08-484-106-22
21	105	4.7	430	1	US-08-484-106-22
22	103	4.6	292	4	US-09-328-352-7538
23	102	4.5	346	2	US-08-602-359A-34
24	101.5	4.5	312	4	US-09-107-532A-5684
25	98.5	4.4	388	1	US-08-232-519-2
26	98.5	4.4	388	1	US-08-456-956-2
27	96.5	4.3	287	4	US-09-584-568C-4

ALIGNMENTS

RESULT 1

US-09-820-001-4
; Sequence 4, Application US/09820001
; Patent No. 6387680
; GENERAL INFORMATION:
; APPLICANT: MERKULOV, Gennady et al
; TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES
; FILE REFERENCE: CL001186
; CURRENT APPLICATION NUMBER: US/09/820,001
; CURRENT FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Human
US-09-820-001-4

Query Match 57.1%; Score 1282; DB 4; Length 392;
Best Local Similarity 63.8%; Pred. No. 6.5e-127; Mismatches 80; Indels 0; Gaps 0;
Matches 234; Conservative 53; Mismatches 80; Indels 0; Gaps 0;

QY	40	AVDPEAFNMISIIHQGYPCPEEYEVATEDGYILSVNRIPIRGLVQPKTKGSRPVVLQHG	99
Db	26	AVDPETNMNVSEIIISYWGFPSEYLVETEDGYILCLNRIPIHGRKNHSDKGPVVFVLQHG	85
QY	100	LVGGSNWTISNLPNNLSGLFILADAGFDVWNGSRGNAMSRKHTLSIDODEFWAFSYDEM	159
Db	86	LIADSSNWTNLSNLSGLFILADAGFDVWNGSRGNAMSRKHTLSIDODEFWAFSYDEM	145
QY	160	ARFDPFAVINFLQTKGQEKIYVGYSGQTGMGTFAPSTMPBLAQIKWYFALAPIATVK	219
Db	146	AKYDLPASINFILNKTGQEVVYVGHSGQTGTFIATFQIPELAKIRKFPALGPVASVA	205
QY	220	HAKSGPTKELLIPDMMIKGLFGKGFVLTQTRFLRQVILYLCQVILDOICSNIMLLGCF	279
Db	206	FCTSPMAKLGRLFDHLIKDLFGDKFELFQSAFLKWLGHVCTHVLKELCGNLCLGCF	265
QY	280	NTNNMMSRASVYAHTLAGTSVQNLHWSQVNSGELRAFQWSETKNLEKCNQPTVR	339
Db	266	NERNLNMSRDVYTHSPAGTSVQNLHWSQVNSGELRAFQWSETKNLEKCNQPTVR	325
QY	340	YVRDMVTFTAWTGGQDMLSNPDKVLLSEVTLNLIYHKNIPFWAHVDFWGLDAPHRM	399
Db	326	YVVKDMLVPTAVWSGHDWLDVVDVNIILLTQITNLVTHESIPWEHLDIFWGLDAPWRL	385
QY	400	YNEIHL 406	

Db 386 YNKIINL 392

RESULT 2

US-08-227-108-17
 ; Sequence 17, Application US/08227108
 ; Patent No. 5807726
 ; GENERAL INFORMATION:
 ; APPLICANT: Blanchard, Claire
 ; APPLICANT: Benicourt, Claude
 ; APPLICANT: Junien, Jean-Louis
 ; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/227,108
 ; FILING DATE: 03-APR-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fanucci, Allan A.
 ; REGISTRATION NUMBER: 30,256
 ; REFERENCE/DOCKET NUMBER: 7620-033
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212 790-9090
 ; TELEFAX: 212 869-8864/9741
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 17:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 377 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-227-108-17

Query Match 51.7%; Score 1161; DB 1; Length 377;
 Best Local Similarity 57.7%; Pred. No. 3.8e-114;
 Matches 213; Conservative 63; Mismatches 93; Indels 0; Gaps 0;

Qy	42	DPEAFNMSIEIIHQGYPCIEYEVATEDGYILSVNRPGLVQPKTKGSRPVVLLQGLV	101
Db	9	NPEANNISOMITWGYPCQIEYEVATEDGYILGVYRPHGKNSENIGKRPVVYLQGLI	68
Qy	102	GGASNWSINLPNNSLGFILADAGFDVWGNRGNASRKHTLSIDQDEFWAFSYDEMAR	161
Db	69	ASAINWIANLPNNSLAFMLADAGFDVWGNRGNATWSRNKVVYSPDSVFWAFSDEMAM	128
Qy	162	FDLPAVINFLQTKGQEKIYVGYSGQTTMGFTAFSTMPDLAKIKRYFALAPIATVKA	221
Db	129	YDLPATINFINVQTKGQEKIHYVGHSGQTTIGTIFAPSTNPLAKKIKTFYALAPVATVKY	188
Qy	222	KSPCTKELLIPDMMIKGLFGKKEFLYQTRFLRQLVYLQGVILDOICSNIMLLGGFNT	281
Db	189	QSPKIKISFPTFLFKLMFGKMFPHYTPDDFLGTGTEVCSREVLDLLCSNTLIFCGFDK	248
Qy	282	NNMMSRASYAAHTLAGTSVQNLHWSQAVNSGELRAFDWGSSETKNLEKCNQPTPVRYR	341
Db	249	KULNVSRFDVYLGHNPAGTSVQDFLHWAQLVRSGKFAEWGSPSQNMLHYNKTTPPYD	308
Qy	342	VRDWTVTAMWTGQDMLSNPQVDMLLSEVTLNLIYHKNIPWAHVDFTWGLDAPERMKN	401
Db	309	VSAMTVEPVAVWNGNDLTADQDQVAMLLPKLSNLLFHKELAYNHLDFIWMADPAQEVYN	368

Qy 402 EIIHMQOE 410
 Db 369 EMISMAED 377

RESULT 3

US-09-073-674-17
 ; Sequence 17, Application US/09073674
 ; Patent No. 5998189
 ; GENERAL INFORMATION:
 ; APPLICANT: Blanchard, Claire
 ; APPLICANT: Benicourt, Claude
 ; APPLICANT: Junien, Jean-Louis
 ; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Warner-Lambert Company
 ; STREET: 2800 Plymouth Road
 ; CITY: Ann Arbor
 ; STATE: Michigan
 ; COUNTRY: U.S.A.
 ; ZIP: 48105
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/073,674
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Crissey, Todd M.
 ; REGISTRATION NUMBER: 37,807
 ; REFERENCE/DOCKET NUMBER: 5072-D1-66-TMC
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 734 622-7530
 ; TELEFAX: 734 622-1553
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 17:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 377 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-073-674-17

Query Match 51.7%; Score 1161; DB 2; Length 377;
 Best Local Similarity 57.7%; Pred. No. 3.8e-114;
 Matches 213; Conservative 63; Mismatches 93; Indels 0; Gaps 0;

Qy	42	DPEAFNMSIEIIHQGYPCIEYEVATEDGYILSVNRPGLVQPKTKGSRPVVLLQGLV	101
Db	9	NPEANNISOMITWGYPCQIEYEVATEDGYILGVYRPHGKNSENIGKRPVVYLQGLI	68
Qy	102	GGASNWSINLPNNSLGFILADAGFDVWGNRGNASRKHTLSIDQDEFWAFSYDEMAM	161
Db	69	ASATWIANLPNNSLAFMLADAGFDVWGNRGNATWSRNKVVYSPDSVFWAFSDEMAM	128
Qy	162	FDLPAVINFLQTKGQEKIYVGYSGQTTMGFTAFSTMPDLAKIKRYFALAPIATVKA	221
Db	129	YDLPATINFINVQTKGQEKIHYVGHSGQTTIGTIFAPSTNPLAKKIKTFYALAPVATVKY	188
Qy	222	KSPCTKELLIPDMMIKGLFGKKEFLYQTRFLRQLVYLQGVILDOICSNIMLLGGFNT	281
Db	189	QSPKIKISFPTFLFKLMFGKMFPHYTPDDFLGTGTEVCSREVLDLLCSNTLIFCGFDK	248
Qy	282	NNMMSRASYAAHTLAGTSVQNLHWSQAVNSGELRAFDWGSSETKNLEKCNQPTPVRYR	341
Db	249	KULNVSRFDVYLGHNPAGTSVQDFLHWAQLVRSGKFAEWGSPSQNMLHYNKTTPPYD	308

[illegible]

APPLICANT: Blanchard, Claire
APPLICANT: Benicourt, Claude
APPLICANT: Junien, Jean-Louis
TITLE OF INVENTION: Recombinant Dog Gastric Lipase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: 03-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fanucci, Allan A.
REGISTRATION NUMBER: 30,256
REFERENCE/DOCKET NUMBER: 7620-033
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-227-108-5

Query Match 50.8%; Score 1141.5; DB 1; Length 380;
Best Local Similarity 56.5%; Pred. No. 4.4e-112;
Matches 210; Conservative 67; Mismatches 90; Indels 5; Gaps 3;
QY 37 PTKAVDPEAFMNISETIIOHGYPCCEYEYVATEDGYILSVNRIPRGLVQPKTGRPVLL 96
DB 8 PT---NPEVTMNIQSMTYWGYPABEYEVVTEGYILGIDRIPYGRKNSENIGRRPVAF 64
QY 97 QHGLVGASNIWNLNPNNSLGFILADAGPDVWNGSRGNAGSKHKTLSIDQDEFWAFSY 156
DB 65 QHGLASATNIWNLNPNNSLAFILADAGYDVLGNSRGNTWARRNLYSPDSVEFWAFSF 124
QY 157 DEMARFDLPVAVINFILQKTQEKIYVYVSGQTTMGFIAPSTMPKLAQIKMYFALAPIA 216
DB 125 DEMAKYDLPATIDFILKKTQDQKLVYVHSGQTTIGFIAPSTNPKLAKRIKTFYALAPVA 184
QY 217 TVKHAKSPGTFKLLPDMIKGLFGKEFLYQTRFLRQ-LVIYLCQGVILDOICSNIMLL 275
DB 185 TVKTEFLNKLMLVPSFLFKLIFGNKIF-YPHFFDQFLATEVCSETRVLDLCSNALFI 243
QY 276 LGGFNTNNMNSRASVYAHTLAGTSVQNTILHWSQAVNSGELAFDWSGSETKNLEKNQ 335
DB 244 ICGFDTNMLNSRLDVLVSHNPAGTSVQNVLHWSQAVKSGKFOAFDWSGSPVQNMHHQS 303
QY 336 TPVRYRVRTMTVPTAMWTGGQDMLSNPDEYKMLSEVTNLIYHKNIPWAHVDFTWGLDA 395
DB 304 MPYYNTDMHVPVIAVWNGNDLLADPHDVLKLLSKLPNLIYHKKIPPNHLDFTWAMDA 363
QY 396 PHRYNEIHL 407
DB 364 QAVYNEIVSM 375

RESULT 11

US-09-073-674-5

; Sequence 5, Application US/09073674

Patent No. 5998189
GENERAL INFORMATION:
APPLICANT: Blanchard, Claire
APPLICANT: Benicourt, Claude
APPLICANT: Junien, Jean-Louis
TITLE OF INVENTION: Recombinant Dog Gastric Lipase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Warner-Lambert Company
STREET: 2800 Plymouth Road
CITY: Ann Arbor
STATE: Michigan
COUNTRY: U.S.A.
ZIP: 48105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: 09/073,674
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Crissey, Todd M.
REGISTRATION NUMBER: 37,807
REFERENCE/DOCKET NUMBER: 5072-D1-66-TMC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 734 622-7530
TELEFAX: 734 622-1553
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-073-674-5

Query Match 50.8%; Score 1141.5; DB 2; Length 380;
Best Local Similarity 56.5%; Pred. No. 4.4e-112;
Matches 210; Conservative 67; Mismatches 90; Indels 5; Gaps 3;
QY 37 PTKAVDPEAFMNISETIIOHGYPCCEYEYVATEDGYILSVNRIPRGLVQPKTGRPVLL 96
DB 8 PT---NPEVTMNIQSMTYWGYPABEYEVVTEGYILGIDRIPYGRKNSENIGRRPVAF 64
QY 97 QHGLVGASNIWNLNPNNSLGFILADAGPDVWNGSRGNAGSKHKTLSIDQDEFWAFSY 156
DB 65 QHGLASATNIWNLNPNNSLAFILADAGYDVLGNSRGNTWARRNLYSPDSVEFWAFSF 124
QY 157 DEMARFDLPVAVINFILQKTQEKIYVYVSGQTTMGFIAPSTMPKLAQIKMYFALAPIA 216
DB 125 DEMAKYDLPATIDFILKKTQDQKLVYVHSGQTTIGFIAPSTNPKLAKRIKTFYALAPVA 184
QY 217 TVKHAKSPGTFKLLPDMIKGLFGKEFLYQTRFLRQ-LVIYLCQGVILDOICSNIMLL 275
DB 185 TVKTEFLNKLMLVPSFLFKLIFGNKIF-YPHFFDQFLATEVCSETRVLDLCSNALFI 243
QY 276 LGGFNTNNMNSRASVYAHTLAGTSVQNTILHWSQAVNSGELAFDWSGSETKNLEKNQ 335
DB 244 ICGFDTNMLNSRLDVLVSHNPAGTSVQNVLHWSQAVKSGKFOAFDWSGSPVQNMHHQS 303
QY 336 TPVRYRVRTMTVPTAMWTGGQDMLSNPDEYKMLSEVTNLIYHKNIPWAHVDFTWGLDA 395
DB 304 MPYYNTDMHVPVIAVWNGNDLLADPHDVLKLLSKLPNLIYHKKIPPNHLDFTWAMDA 363
QY 396 PHRYNEIHL 407
DB 364 QAVYNEIVSM 375

RESULT 12

US-09-186-489-2
 ; Sequence 2, Application US/09186489
 ; Patent No. 6375947
 ; GENERAL INFORMATION:
 ; APPLICANT: Bolen, Paul L.
 ; APPLICANT: Chak, Paul L.
 ; APPLICANT: Scharpf Jr., Lewis G.
 ; TITLE OF INVENTION: Purified Recombinant Kid Pregastric Esterase, and
 ; TITLE OF INVENTION: Processes for its Production and Use
 ; FILE REFERENCE: 5499/3
 ; CURRENT APPLICATION NUMBER: US/09/186,489
 ; CURRENT FILING DATE: 1998-11-05
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: Patent in Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 378
 ; TYPE: PRT
 ; ORGANISM: Kid (Goat)
 ; US-09-186-489-2

Query Match 50.2%; Score 1127; DB 4; Length 378;
 Best Local Similarity 55.1%; Pred. No. 1.5e-110;
 Matches 205; Conservative 66; Mismatches 101; Indels 0; Gaps 0;
 QY 40 AVDEAFPMNISEIIHQGYPCSEYEVATEDGYILSVNRIPRGLVQPKTKGSRPVVLLQHG 99
 Db 6 AKNPEASMNVSQMSIFWGYPSMKHVTADGYILQVYRPHGKNDANHLGQRPVFLQHG 65
 QY 100 LVGCASWISNLPNNSLGFILADAGFDVWNGSRGNARSKKHTLSIDODEFWAFSDYM 159
 Db 66 LLASATNWSNLPNNSLGFILADAGFDVWNGSRGNARSKKHTLSIDODEFWAFSDYM 125
 QY 160 ARFDPAVINFIQKQGOEKIYVGVYSGQTTMGFTAFSTMPPELAQIKMYFALAPATVK 219
 Db 126 AEYDLSTIDFILKRTQCKLHVGHSGQTTIGFVAFSTNPTLAEKIEVPHALAPATVK 185
 QY 220 HAKSPGTFKLLPDMIMKGLFGKKEFLYQTRFRLQVLYLGGVILDOICSNMILLGGF 279
 Db 186 HTQSLFNKALIPHLFKIIFGNKMFYPHNFQFLGVEVCSRETLDVLCKNALFAITGA 245
 QY 280 NTNNMNSRASVYAAHTLAGTSVONILHWSQVNSGELRAFQWGETKLEKNCQPTVR 339
 Db 246 DNKNFNSRLDYYVAHPAGASVQNLHWRQAIKSKGFOAFDNGASVENLMHYNQTPPI 305
 QY 340 YVRDVTPTAMTGGQDMLSNPDEYKLLSEVTNLIYHKNIPWAHVDPIWGLDAPHRM 399
 Db 306 YNLTMNVPIAVNSAGQDLADPDQVDLLSKLSNLHKEIPNYNHLDFIWMADAPQEV 365
 QY 400 YNEIHLMOQEE 411
 Db 366 YNEIISLMKDK 377

RESULT 13
 US-10-043-665B-2
 ; Sequence 2, Application US/10043665B
 ; Patent No. 6582948
 ; GENERAL INFORMATION:
 ; APPLICANT: Bolen, Paul L.
 ; APPLICANT: Chak, Paul L.
 ; APPLICANT: Scharpf Jr., Lewis G.
 ; TITLE OF INVENTION: Recombinant Kid Pregastric Esterase and Methods for Its
 ; TITLE OF INVENTION: Production and Use
 ; FILE REFERENCE: IFF-0009
 ; CURRENT APPLICATION NUMBER: US/10/043,665B
 ; CURRENT FILING DATE: 2002-09-24
 ; PRIOR APPLICATION NUMBER: US 09/186,489
 ; PRIOR FILING DATE: 1998-11-05
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: Patent in Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 378
 ; TYPE: PRT

; ORGANISM: Kid (Goat)
 US-10-043-665B-2
 Query Match 50.2%; Score 1127; DB 4; Length 378;
 Best Local Similarity 55.1%; Pred. No. 1.5e-110;
 Matches 205; Conservative 66; Mismatches 101; Indels 0; Gaps 0;
 QY 40 AVDEAFPMNISEIIHQGYPCSEYEVATEDGYILSVNRIPRGLVQPKTKGSRPVVLLQHG 99
 Db 6 AKNPEASMNVSQMSIFWGYPSMKHVTADGYILQVYRPHGKNDANHLGQRPVFLQHG 65
 QY 100 LVGCASWISNLPNNSLGFILADAGFDVWNGSRGNARSKKHTLSIDODEFWAFSDYM 159
 Db 66 LLASATNWSNLPNNSLGFILADAGFDVWNGSRGNARSKKHTLSIDODEFWAFSDYM 125
 QY 160 ARFDPAVINFIQKQGOEKIYVGVYSGQTTMGFTAFSTMPPELAQIKMYFALAPATVK 219
 Db 126 AEYDLSTIDFILKRTQCKLHVGHSGQTTIGFVAFSTNPTLAEKIEVPHALAPATVK 185
 QY 220 HAKSPGTFKLLPDMIMKGLFGKKEFLYQTRFRLQVLYLGGVILDOICSNMILLGGF 279
 Db 186 HTQSLFNKALIPHLFKIIFGNKMFYPHNFQFLGVEVCSRETLDVLCKNALFAITGA 245
 QY 280 NTNNMNSRASVYAAHTLAGTSVONILHWSQVNSGELRAFQWGETKLEKNCQPTVR 339
 Db 246 DNKNFNSRLDYYVAHPAGASVQNLHWRQAIKSKGFOAFDNGASVENLMHYNQTPPI 305
 QY 340 YVRDVTPTAMTGGQDMLSNPDEYKLLSEVTNLIYHKNIPWAHVDPIWGLDAPHRM 399
 Db 306 YNLTMNVPIAVNSAGQDLADPDQVDLLSKLSNLHKEIPNYNHLDFIWMADAPQEV 365
 QY 400 YNEIHLMOQEE 411
 Db 366 YNEIISLMKDK 377

RESULT 14
 US-09-348-930A-9
 ; Sequence 9, Application US/09348930A
 ; Patent No. 6573431
 ; GENERAL INFORMATION:
 ; APPLICANT: Lence, P
 ; APPLICANT: Gruber, V
 ; APPLICANT: Baudino, S
 ; APPLICANT: Merot, B
 ; APPLICANT: Benicourt, C
 ; APPLICANT: Cudrey, C
 ; TITLE OF INVENTION: Recombinant Preducodenal Lipases and Polypeptides Derivatives Prod
 ; TITLE OF INVENTION: Plants, Processes for Obtaining Them and Their Uses
 ; FILE REFERENCE: 18433/2012
 ; CURRENT APPLICATION NUMBER: US/09/348,930A
 ; CURRENT FILING DATE: 1999-07-02
 ; PRIOR APPLICATION NUMBER: 08/945,321
 ; PRIOR FILING DATE: 1998-02-12
 ; PRIOR APPLICATION NUMBER: PCT/FR96/00606
 ; PRIOR FILING DATE: 1996-03-19
 ; NUMBER OF SEQ ID NOS: 37
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 9
 ; LENGTH: 398
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-348-930A-9

Query Match 49.8%; Score 1119; DB 4; Length 398;
 Best Local Similarity 52.2%; Pred. No. 1.1e-109;
 Matches 210; Conservative 73; Mismatches 107; Indels 12; Gaps 2;
 QY 17 MWLLILVAVYFQNVNSVEMPTKAV-----DPEAFMNISETIIOHGYPCSEYEVATED 69
 Db 1 MWLLITWASLI-----SVLGTTHGLFGKLGHPGSEVTMNIQSMTIWTGTFNBEIEYVTTED 55
 QY 70 GYILSVNRIPRGLVQPKTKGSRPVVLLQHGVLGASWISNLPNNSLGFILADAGFDVW 129

Db 56 GYILEVNRIPYKKNKSGNTQRPVFLQHGLLASATNWSLNPNSLAFILADAGYDVL 115
QY 130 GNSRGNAWSKHKHTLSDDQEFWAFSDENAREDLPAVINFILOKTOGQEKIYVGYSGQT 189
Db 116 GNSRGNTWARNLYYSPDSVEFWAFSDENAREDLPAVINFILOKTOGQEKIYVGYSGQT 175
QY 190 TMGFIAFSTMPLELAQKIMYFALAPIATVHAKSPGTFKLLPDMIMKGLFGKGEFLYOT 249
Db 176 TIGFIAFSTNPSLAKRIKTFYALAPVATVYTKSLINKLRFVPSLKFIFGDKIFYPHN 235
QY 250 RELRQVILYLCGGVILDOICSNIMLLGGENTNNMNSRASVYAAHTLAGTSVONILHWS 309
Db 236 FFOQFLATEVCSREMLNLLCSNALFIICGPDSCNFTSLDVLVLSHPAGTSVQNNFHT 295
QY 310 QAVNSGELRAFDRGSETKNLEKNCQTFVRYRVRDMTVPMTAMTGGQDMLSNPDEYKMLL 369
Db 296 QAVKSGKGFQAYDNGSPVQNRMHYDQSQPPYVNTAMNVPVAVNGGKDLADPDQVGLLL 355
QY 370 SEVNTLIYHKNIPENAHVDFIWLGDAPHRMNIEIHLMOQEE 411
Db 356 PKLPNLIIYHKEIPFYNHLDPIWADAPQEVYNDIVSMISEDK 397

RESULT 15

US-08-227-108-16
; Sequence 16 Application US/08227108
; Patent No. 5807726
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/227,108
; FILING DATE: 03-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Panucci, Allan A.
; REGISTRATION NUMBER: 30,256
; REFERENCE/DOCKET NUMBER: 7620-033
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-227-108-16

Query Match 48.9% Score 1098.5; DB 1; Length 380;
Best Local Similarity 54.6%; Pred. No. 1.5e-107;
Matches 202; Conservative 69; Mismatches 98; Indels 1; Gaps 1;
QY 43 PEAPMNISETIQHGVCPEYEVATEDGYILSVNRIIPRGVQPKTKGSRPVLLOHGLVG 102

Db 10 PEVTMNIQSMTITYMGYPNEEYEVVTDGYILEVNRIPYKKNKSGNTQRPVFLQHGLLA 69
QY 103 GASNWSLNPNSLGFILADAGEDVWNGSRGNASRKHKHTLSIDQDEFW-AFSYDEMAR 161
Db 70 SATNWSLNPNSLAFILADAGEDVWNGSRGNWNRNLYYSPDSVEFWAFSDEMAK 129
QY 162 FDLPAVINFILOKTOGQEKIYVGYSGQTGTFIAFSTMPLELAQKIMYFALAPIATVHKA 221
Db 130 YDLPATIDFIVKKTGQKQLHVGHSGQTGTFIAFSTNPSLAKRIKTFYALAPVATVKYT 189
QY 222 KSPGTFKLLLPDMIMKGLFGKGEFLYOTRFLRQVILYLCGGVILDOICSNIMLLGGENT 281
Db 190 KSLINKLRFVPSLKFIFGDKIFYPHNFPDQFLATEVCSREMLNLLCSNALFIICGDS 249
QY 282 NNMNSRASVYAAHTLAGTSVONILHWSQAVNSGELRAFDRGSETKNLEKNCQTFVRYR 341
Db 250 KNFTSLRLDVLVLSHPAGTSVQNNFHTQAVKSGKGFQAYDNGSPVQNRMHYDQSQPPYV 309
QY 342 VRDMTVPMTAMTGGQDMLSNPDEYKMLLSEVNTLIYHKNIPENAHVDFIWLGDAPHRMYN 401
Db 310 VTAMNVPVAVNGGKDLADPDQVGLLLPKLPNLIIYHKEIPFYNHLDPIWADAPQEVYN 369
QY 402 EIIHLMOQEE 411
Db 370 DIVSMISEDK 379

Search completed: February 19, 2004, 14:23:46
Job time : 16.5683 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2004, 14:16:19 ; Search time 9.11439 Seconds
(without alignments)
2012.248 Million cell updates/sec

Title: US-10-042-431-49

Perfect score: 2076

Sequence: 1 VHMPTKAYDPEAFNMISII.....IHLNQBEETNLQGRCEAVL 390

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1289	62.1	399	1	LICH HUMAN
2	1220.5	58.8	397	1	LICH RAT
3	1194.5	57.5	397	1	LICH MOUSE
4	1161	55.9	395	1	LIPG RAT
5	1141.5	55.0	398	1	LIPG CANFA
6	1125	54.2	397	1	LIPG BOVIN
7	1109	53.4	398	1	LIPG HUMAN
8	663	31.9	394	1	LIP3 DROME
9	546	26.3	439	1	LIP1 DROME
10	422.5	20.4	548	1	TGL1 YEAST
11	111.5	5.4	987	1	YD94 METUA
12	105	5.1	430	1	ORC2 CABEL
13	98.5	4.7	436	1	PAPA CAVPO
14	96.5	4.6	287	1	Y193 HAEIN
15	93.5	4.5	278	1	PRXC STRAU
16	93	4.5	569	1	PRXD STRAU
17	92	4.4	748	1	PTIP ECOLI
18	92	4.4	748	1	PTIP SALTY
19	91	4.4	370	1	P2C2 SCHPO
20	90	4.3	277	1	BPA2 STRAU
21	89	4.3	330	1	EXOA RHIME
22	88.5	4.3	185	1	Y2H BACSU
23	88	4.2	1259	1	LINI HUMAN
24	87	4.2	473	1	SCRB LACIA
25	86.5	4.2	382	1	YDHF HAEIN
26	86.5	4.2	685	1	INVA SALTY
27	86.5	4.2	869	1	CPAC ECOLI
28	86.5	4.2	1122	1	Y3C3 YEAST
29	86	4.1	422	1	Y3B3 SHEON
30	86	4.1	1866	1	VGNB CPMV
31	85.5	4.1	428	1	Y563 SCHPO
32	85	4.1	292	1	PAEB NEUCR
33	85	4.1	429	1	ARB2_ECOLI

RESULT 1

ID	LICH HUMAN	STANDARD;	PRT;	399 AA.
AC	P38571; Q16529; Q96EJO;			
DT	01-OCT-1994 (Rel. 30, Created)			
DT	01-OCT-1994 (Rel. 30, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Lysosomal acid lipase/cholesterol ester hydrolase precursor			
DE	(EC 3.1.1.13) (LAL) (Acid cholesterol ester hydrolase) (Sterol			
DE	esterase) (Lipase A) (Cholesteryl esterase).			
GN	LIPA.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 196-212; 277-297 AND 305-315.			
RX	MEDLINE=92042192; PubMed=1718995;			
RA	Anderson R.A.; Sando G.N.;			
RT	"Cloning and expression of cDNA encoding human lysosomal acid			
RT	lipase/cholesterol ester hydrolase. Similarities to gastric and			
RT	lingual lipases.";			
RL	J. Biol. Chem. 266:22479-22484 (1991).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	TISSUE=Liver;			
RX	MEDLINE=94155897; PubMed=8112342;			
RA	Ameis D., Merkel M., Eckerskorn C., Greten H.;			
RT	"Purification, characterization and molecular cloning of human			
RT	hepatic lysosomal acid lipase.";			
RL	Eur. J. Biochem. 219:905-914 (1994).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	TISSUE=Liver;			
RX	MEDLINE=96363957; PubMed=8725147;			
RA	Du H., Witte D.P., Grabowski G.A.;			
RT	"Tissue and cellular specific expression of murine lysosomal acid			
RT	lipase mRNA and protein.";			
RL	J. Lipid Res. 37:937-949 (1996).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	TISSUE=Placenta;			
RX	MEDLINE=22388257; PubMed=12477932;			
RA	Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Strausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Abschul S.F., Zeeb B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marushina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,			
RA	Brownstein W.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahey J., Helton E., Kettenan M., Madan A., Rodriguez S., Sanchez A.,			
RA	Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			

ALIGNMENTS

Q89ak4 buchnera ap
Q9uuJ9 schizosacch
P97313 mus musculus
P38295 saccharomyc
P36430 bacillus su
P52408 prunus pers
Q9QYF9 mus musculus
P55882 salmonella
Q31158 pseudomonas
P9KVB1 vibrio chol
Q04635 staphylococ
P16115 thermotoga

34 85 4.1 429 1 BIOA BUCBP
35 85 4.1 521 1 ACHI SCHPO
36 85 4.1 4128 1 PRXD MOUSE
37 84 4.0 451 1 YB27 YEAST
38 84 4.0 804 1 SYL BACSU
39 83.5 4.0 350 1 E13B PRUPE
40 83.5 4.0 375 1 NDR3 MOUSE
41 83 4.0 266 1 THID SALTY
42 83 4.0 273 1 PRXC PSEPL
43 83 4.0 429 1 PUR2 VIBCH
44 83 4.0 641 1 LIP STAHY
45 82.5 4.0 319 1 LDH THEMA

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
 RA Schnerch A., Schein J.E., Jones S.U.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [5]
 RP VARIANT CESD/WD PRO-200.
 RX MEDLINE-94195814; PubMed-8146180;
 RA Anderson R.A., Byrum R.S., Coates P.M., Sando G.N.;
 RT "Mutations at the lysosomal acid cholesteryl ester hydrolase gene
 RT locus in Wolman disease.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:2718-2722(1994).
 RN [6]
 RP VARIANTS CESD ARG-129 AND PRO-129.
 RX MEDLINE-98235576; PubMed-9633819;
 RA Ries S., Buechler C., Schindler G., Alanidis C., Aneis D., Gasche C.,
 RA Jung N., Schambach A., Fehringer P., Vanier M.T., Belli D.C.,
 RA Greten H., Schmitz G.;
 RT "Differant missense mutations in histidine-108 of lysosomal acid
 RT lipase cause cholesteryl ester storage disease in unrelated compound
 RT heterozygous and hemizygous individuals.";
 RL Hum. Mutat. 12:44-51(1998).
 CC -1- FUNCTION: CRUCIAL FOR THE INTRACELLULAR HYDROLYSIS OF CHOLESTERYL
 CC ESTERS AND TRIGLYCERIDES THAT HAVE BEEN INTERNALIZED VIA RECEPTOR-
 CC MEDIATED ENDOCYTOSIS OF LIPOPROTEIN PARTICLES. IMPORTANT IN
 CC MEDIATING THE EFFECT OF LDL (LOW DENSITY LIPOPROTEIN) UPTAKE ON
 CC SUPPRESSION OF HYDROXYMETHYLGLUTARYL-COA REDUCTASE AND ACTIVATION
 CC OF ENDOGENOUS CELLULAR CHOLESTERYL ESTER FORMATION.
 CC -1- CATALYTIC ACTIVITY: A sterol ester + H(2)O = a sterol + a fatty
 CC acid
 CC -1- SUBCELLULAR LOCATION: Lysosomal.
 CC -1- DISEASE: DEFECTS IN LIPA ARE THE CAUSE OF THE SEVERE INFANTILE-
 CC ONSET WOLMAN DISEASE (WD) AND THE Milder LATE-ONSET CHOLESTERYL
 CC ESTER STORAGE DISEASE (CESD).
 CC -1- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; M74775; AAA59519.1; -;
 DR EMBL; U04285; AAB60327.1; -;
 DR EMBL; U04286; AAB60327.1; JOINED.
 DR EMBL; U04287; AAB60327.1; JOINED.
 DR EMBL; U04288; AAB60327.1; JOINED.
 DR EMBL; U04290; AAB60327.1; JOINED.
 DR EMBL; U04291; AAB60327.1; JOINED.
 DR EMBL; U04292; AAB60327.1; JOINED.
 DR EMBL; U04293; AAB60327.1; JOINED.
 DR EMBL; X76488; CAA54026.1; -;
 DR EMBL; X31690; CAA83495.1; -;
 DR EMBL; U08464; AAB60328.1; -;
 DR EMBL; BC012287; AAL12287.1; -;
 DR PIR; S41408; S41408.
 DR Genew; HGNC:6617; LIPA.
 DR MIM; 278000; -;
 DR GO; GO:0005764; C:lysosome; TAS.
 DR GO; GO:0006487; P:N-linked glycosylation; TAS.
 DR InterPro; IPR000073; A/b hydrolase.
 DR InterPro; IPR000734; Lipase.
 DR InterPro; IPR000379; Ser esterase site.
 DR Pfam; PF04083; abhydro_lipase_1.
 DR Pfam; PF00561; abhydro_lipase; 1.
 DR PROSITE; PS00120; LIPASE_SER; 1.
 KW Hydrolase; lipid degradation; Glycoprotein; Signal; Lysosome;
 KW Disease mutation; Polymorphism.
 FT SIGNAL 1 21
 FT POTENTIAL.

FT CHAIN 22 399 LYSOSOMAL ACID LIPASE/CHOLESTERYL ESTER
 FT ACT_SITE 174 174 HYDROLASE.
 FT ACT_SITE 374 374 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT CARBOHYD 36 36 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT CARBOHYD 72 72 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 101 101 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 161 161 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 273 273 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 321 321 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT VARIANT 16 16 P -> T.
 FT VARIANT 129 129 H -> P (IN CESD).
 FT VARIANT 129 129 H -> R (IN CESD).
 FT VARIANT 200 200 L -> P (IN CESD AND WD).
 FT CONFLICT 23 23 G -> R (IN REF. 3).
 FT CONFLICT 29 29 V -> L (IN REF. 4).
 SQ SEQUENCE 399 AA; 45415 MW; 55F27391306B609A CRC64;
 Query Match 62.1%; Score 1289; DB 1; Length 399;
 Best Local Similarity 63.5%; Pred. No. 7.6e-107;
 Matches 235; Conservative 55; Mismatches 80; Indels 0; Gaps 0;
 QY 7 AVDEAFNNISEIIHQGYPCIEYEYATEDGYILSVNRIPIRGVLQPKTGSRRVYLQHG 66
 DB 28 AVDEPTNNVSEIISYWGFPSEYLVETEDGYILCLNRIPIRGVNSDKGPKVFLQHG 87
 QY 67 LVGASWISNLPNNSLGFILADAGFDVVMGNSRGNASRKHKTLSIDQDFWAFSYDEM 126
 DB 88 LLDSSNNVNTLANSLSLGFILADAGFDVVMGNSRGNATSKHKTLSVSDQDFWAFSYDEM 147
 QY 127 ARFDLPVINFILQKQYVGVSGQTMGFIAFSTMPPELAQIKTIFALAPIATVK 186
 DB 148 AKYDLPASINFLNKTQEQVYVGHSGQTTIGFIAPSQIPELAKIKMFALGPVASVA 207
 QY 187 HAKSPGKFKLLPDMMLKGLFGKKEFLYQTRQLQVLYLGGVILDCGVDICSNIMLLGGF 246
 DB 208 FCTSPMAKGLRPLDHLKDLFGDKFLPQSAFLKWLGVHVTHTVILKELGNCJFLGCP 267
 QY 247 NTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNSGELRAFDMGSETKNEKCNQPTFVR 306
 DB 268 NERNLNSRVDVYTHSPAGTSVQNLHWSQAVKQFQAFDMGSSAKNYFHYNQSYPT 327
 QY 307 YRVDMVTPTAMWTGGODWLSNPEDVZMLSEVNTLYHKNIPEWAHVDEINGLDAPHRM 366
 DB 328 YNVKMLVPTAVWSGGHDLADVDVYNILLTQITNLVFNHESIPWEHLDFINGLDAPWRL 387
 QY 367 YNEIHLMOQ 376
 DB 388 YNKIINLWK 397
 RESULT 2
 LICH RAT
 ID LICH RAT STANDARD; PRT; 397 AA.
 AC Q64194;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Lysosomal acid lipase/cholesteryl ester hydrolase precursor
 DE (EC 3.1.1.13) (LAL) (Acid cholesteryl ester hydrolase) (Sterol
 DE esterase) (Lipase A) (Cholesteryl esterase).
 GN LIPA OR LAL.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;

RX MEDLINE=96129534; PubMed=8576647;
 RA Nakagawa H., Matsubara S., Kuriyama M., Yoshidome H., Fujiyama J.,
 RA Yoshida H., Osame M.;
 RT "Cloning of rat lysosomal acid lipase cDNA and identification of the
 RT mutation in the rat model of Wolman's diseasee.";
 RL J. Lipid Res. 36:2212-2218 (1995).
 CC -!- FUNCTION: CRUCIAL FOR THE INTRACELLULAR HYDROLYSIS OF CHOLESTERYL
 CC ESTERS AND TRIGLYCERIDES THAT HAVE BEEN INTERNALIZED VIA RECEPTOR-
 CC MEDIATED ENDOCYTOSIS OF LIPOPROTEIN PARTICLES. IMPORTANT IN
 CC MEDIATING THE EFFECT OF LDL (LOW DENSITY LIPOPROTEIN) UPTAKE ON
 CC SUPPRESSION OF HYDROXYMETHYLGLUTARYL-COA REDUCTASE AND ACTIVATION
 CC OF ENDOGENOUS CELLULAR CHOLESTERYL ESTER FORMATION (BY
 CC similarity).
 CC -!- CATALYTIC ACTIVITY: A sterol ester + H(2)O = a sterol + a fatty
 CC acid.
 CC -!- SUBCELLULAR LOCATION: Lysosome.
 CC -!- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; S81497; AAB36043.2; -;
 CC InterPro; IPR000073; A/b hydrolase.
 CC InterPro; IPR000734; Lipase.
 CC InterPro; IPR000379; Ser esters site.
 CC Pfam; PF04083; abhydro lipase; 1.
 CC Pfam; PF00561; abhydro lipase; 1.
 CC PROSITE; PS00120; LIPASE SER; 1.
 CC PROSITE; PS00120; LIPASE SER; 1.
 CC Hydrolase; Lipid degradation; Glycoprotein; Signal; Lysosome.
 CC SIGNAL 1 19
 CC CHAIN 20 397
 CC
 CC ACT SITE 172 172
 CC ACT SITE 372 372
 CC CARBOHYD 34 34
 CC CARBOHYD 99 99
 CC CARBOHYD 159 159
 CC CARBOHYD 271 271
 CC CARBOHYD 319 319
 CC SEQUENCE 397 AA; 45186 MW; 97A38595A0523947 CRC64;
 SQ
 Query Match 58.8%; Score 1220.5; DB 1; Length 397;
 Best Local Similarity 58.2%; Fred. No. 9.1e-101;
 Matches 220; Conservative 67; Mismatches 84; Indels 7; Gaps 3;
 4 FT---KAYDPEAFMNIISIIHQGYPCBEYEVATEDGYLSVNRIPRGLVQPKTGRPV 60
 20 PGTGISAVDPANMNVTEIIMHWGP--EHSVQICDGVILGVHRIPIGRKNQPKPKPV 77
 61 VLLQ--HGLVGASNIENLNFSLGFTIADAGFDVWNGSRGNWAKRKHKLTSIDQDEF 118
 78 VYLQWRHGFADSNVWNTIDNLSLGTIADAGFDVWNGSRGNTWKRKHKLTSVSDQY 137
 119 WAFSDENARDELPAVINFIKQGOEKIYVYGVSQGTMTGFIATFTWPELAQIKMYEA 178
 138 WAFSDENAKYDLPAVINILNKQGOELYNVHSGQCTIGFIATFQSPHKLAKKMYKFFA 197
 179 LAPATVHAKSPGTFKLLLDMMIKGLFGKKEFLYQTRFRLQRIVLQGOVILDQICSN 238
 198 LAPVLSLNFASGPMVKGLRDLLEDFGQKFLQPSAMVKWLSITCTHIVMKELCAN 257
 239 IMLLGGFTNNMNSRSVVAATLACTSVQNTILHWSQVNSGELRAFQWGSSEKLEK 298
 258 IFFLCIGNEKLNLSRVVDVTTCPAGTSVQNVHWTQVVKYHKLQAFDNGSSDKNYFH 317
 299 CNOQTFVRYVRDMVTPTAMVTGGODWLSNPENKMLLSVNTLIYHKNIPEWAVDFIW 358
 318 YNQSYPPPLYSIKDMLPTALWGGKDLADTSDINILLTEIPTLVYHKNIPEWHDLFIW 377

QY 359 GLDAPHRMYNEIHLMOQ 376
 Db 378 GLDAPRWLYNEVSLMKK 395

RESULT 3

LICH MOUSE
 ID LICH MOUSE STANDARD; PRT; 397 AA.
 AC QZOM5;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Lysosomal acid lipase/cholesteryl ester hydrolase precursor
 DE (SC 3.1.1.13) (LAL) (Acid cholesteryl ester hydrolase) (Sterol
 DE esterase) (Lipase A) (Cholesteryl esterase).
 GN LIPA OR LIPI.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6 X CBA; TISSUE=Liver;
 RX MEDLINE=96363957; PubMed=8725147;
 RA Du H., Witte D.P., Grabowski G.A.;
 RT "Tissue and cellular specific expression of murine lysosomal acid
 RT lipase mRNA and protein.";
 RT J. Lipid Res. 37:937-949 (1996).
 RL
 CC -!- FUNCTION: CRUCIAL FOR THE INTRACELLULAR HYDROLYSIS OF CHOLESTERYL
 CC ESTERS AND TRIGLYCERIDES THAT HAVE BEEN INTERNALIZED VIA RECEPTOR-
 CC MEDIATED ENDOCYTOSIS OF LIPOPROTEIN PARTICLES. IMPORTANT IN
 CC MEDIATING THE EFFECT OF LDL (LOW DENSITY LIPOPROTEIN) UPTAKE ON
 CC SUPPRESSION OF HYDROXYMETHYLGLUTARYL-COA REDUCTASE AND ACTIVATION
 CC OF ENDOGENOUS CELLULAR CHOLESTERYL ESTER FORMATION (BY
 CC similarity).
 CC -!- CATALYTIC ACTIVITY: A sterol ester + H(2)O = a sterol + a fatty
 CC acid.
 CC -!- SUBCELLULAR LOCATION: Lysosome.
 CC -!- TISSUE SPECIFICITY: Expressed at low levels in most tissues. High
 CC level expression is found in hepatocytes and splenic and thymic
 CC cells. Very high level expression is observed in cells of the
 CC small intestinal villi, the zona fasciculata and reticularis of
 CC the adrenal cortex, pancreatic acini, and renal tubular
 CC epithelium.
 CC -!- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; Z31689; CAA83494.1; -;
 CC MGD; MGI:96789; Lip1.
 CC InterPro; IPR000073; A/b hydrolase.
 CC InterPro; IPR000734; Lipase.
 CC InterPro; IPR000379; Ser esters site.
 CC Pfam; PF04083; abhydro lipase; 1.
 CC Pfam; PF00561; abhydro lipase; 1.
 CC PROSITE; PS00120; LIPASE SER; 1.
 CC PROSITE; PS00120; LIPASE SER; 1.
 CC Hydrolase; Lipid degradation; Glycoprotein; Signal; Lysosome.
 CC SIGNAL 1 19
 CC CHAIN 20 397
 CC
 CC ACT SITE 172 172
 CC ACT SITE 372 372
 CC CARBOHYD 34 34
 CC CARBOHYD 99 99
 CC CARBOHYD 159 159
 CC CARBOHYD 271 271

FT CARBOHYD 319 319 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 397 AA; 45551 MW; P886C39E1CCFA91P CRC64;
 Query Match 57.5%; Score 1194.5; DB 1; Length 397;
 Best Local Similarity 57.3%; Pred. No. 1.9e-98;
 Matches 216; Conservative 66; Mismatches 92; Indels 3; Gaps 1;
 QY 3 MPT---KAYDPAFNMISIIHQGYPCCEYEVEATEDGYILSVNRIIPRGLVQPKKTGRSP 59
 DB 19 VPTGTVSADVPVNMVNTIIMRWGYPGEHSLVTCGDYILSIHRIIPRGMKHFQKGRP 78
 QY 60 VLLQHLGVGASNTSNIIPNNSLGPILADAGFVMMGNSRGNWSRKHKLISIDQDEFW 119
 DB 79 VVYLQHLGLADSNWNTINDNSSLGFLADRGFDVMMGNSRGNWLSKHKLISVSDQDEFW 138
 QY 120 AFSYDMDAPDLPVAVINFIQKQGYPCCEYEVEATEDGYILSVNRIIPRGLVQPKKTGRSP 179
 DB 139 AFSYDMDAPDLPVAVINFIQKQGYPCCEYEVEATEDGYILSVNRIIPRGLVQPKKTGRSP 198
 QY 180 APIATVKHAKSPOTKFLPDLMMIKGLFGKGFYQTRFLRQVLYLQGVILQICNSI 239
 DB 199 APVLSNFASGFLQLGLPDLPLKDMFGKQFLPQSAWLKWSIHVCTHVMKELCANV 258
 QY 240 MLLGGFNTNMMSRASVYAAHTLACTSVQNLHWSQAVNSGELRADWGSSEKLEKC 299
 DB 259 FFLCGFNEKLMNSRDVITTHCPAELLVQNLHWSQAVNSGELRADWGSSEKLEKC 318
 QY 300 NQPTPVRYRVMTVPTAMWTGQDLSNPEDVYMLSEVTLNLYHKNIPEWAHVDPIWG 359
 DB 319 NQSPFSPYXIKMRLPTALWGGDRDLADINDITILLQIKLVYHKNIPEWHDLDPIWG 378
 QY 360 LDAPHMRYNEIHLMOQ 376
 DB 379 LDAPWKLYDEIISLMKK 395
 RESULT 4
 LIPG RAT
 ID LIPG RAT STANDARD; PRT; 395 AA.
 AC P04634;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Triacylglycerol lipase, lingual precursor (EC 3.1.1.3) (Lingual
 lipase).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85215587; PubMed=3839077;
 RA Doeherty A.J.P., Bodmer M.W., Angal S., Verger R., Riviere C.,
 RA Lowe P.A., Lyons A., Entage J.S., Harris T.J.R.;
 RA "Molecular cloning and nucleotide sequence of rat lingual lipase
 cDNA";
 RL Nucleic Acids Res. 13:1891-1903(1985).
 CC -1- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O = diacylglycerol + a
 fatty acid anion.
 CC -1- MISCELLANEOUS: LINGUAL LIPASE IS SECRETED BY THE SEROUS (VON
 ENER'S) GLANDS AT THE BACK OF THE RAT TONGUE.
 CC -1- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X02309; CAA26179.1; --
 CC EMBL; A01157; CAA00136.1; --

DR PIR: A23045; LIRTT.
 DR InterPro; IPR000073; A/b hydrolase.
 DR InterPro; IPR000734; Lipase.
 DR InterPro; IPR000379; Ser. esters. site.
 DR Pfam; PF00561; abhydro. lipase; 1.
 DR Pfam; PF00561; abhydro. lipase; 1.
 DR PROSITE; PS00120; LIPASE SER; 1.
 KW Hydrolase; Lipid degradation; Glycoprotein; Signal.
 FT SIGNAL 1 18
 FT CHAIN 19 395 TRIACYLGLYCEROL LIPASE. LINGUAL.
 FT ACT_SITE 171 171 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 342 342 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 371 371 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 245 254 BY SIMILARITY.
 FT CARBOHYD 33 33 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 270 270 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 395 AA; 44588 MW; E601854A923522EA CRC64;
 Query Match 55.9%; Score 1161; DB 1; Length 395;
 Best Local Similarity 57.7%; Pred. No. 1.8e-95;
 Matches 213; Conservative 63; Mismatches 93; Indels 0; Gaps 0;
 QY 9 DPAPFMNISEIIHQGYPCCEYEVEATEDGYILSVNRIIPRGLVQPKKTGRSPVLLQGLV 68
 DB 27 NPEANMISQMITWGYPCCEYEVEATEDGYILSVNRIIPRGLVQPKKTGRSPVLLQGLV 86
 QY 69 GGSNWTISNIPNNSLGPILADAGFVMMGNSRGNWSRKHKLISIDQDEFWAFSYDEMAR 128
 DB 87 ASATNWIAPNNSLGPILADAGFVMMGNSRGNWSRKHKLISIDQDEFWAFSYDEMAR 146
 QY 129 FDLPAVINFIQKQGYPCCEYEVEATEDGYILSVNRIIPRGLVQPKKTGRSPVLLQGLV 188
 DB 147 YDLPAVINFIQKQGYPCCEYEVEATEDGYILSVNRIIPRGLVQPKKTGRSPVLLQGLV 206
 QY 189 KSPGTEKLLPDMWIKGLFGKGFYQTRFLRQVLYLQGVILQICNSIIMLLGGFNT 248
 DB 207 QSPLEKLSFIPTFLKLMFGKMFPHYPTDPLGTEVSCREVLDCSNLTFIFGFGDK 266
 QY 249 NNMMSRASVYAAHTLACTSVQNLHWSQAVNSGELRADWGSSEKLEKCNQPTPVRYR 308
 DB 267 KNLNVSFVYLGHNPAQTSVQDFLHWAQLVRSKGFQAFNMGSPSQNMLHYNQKTPPEYD 326
 QY 309 VRDMVTPTAMWTGQDLSNPEDVYMLSEVTLNLYHKNIPEWAHVDPIWGLDAPHRMYN 368
 DB 327 VSAMTVPVAVWNGNDIADPQDVAMLLPKLSNLFHKEILAYNHLDFIWMADAPQEVYN 386
 QY 369 EIIHLMQOE 377
 DB 387 EMISMAED 395
 RESULT 5
 LIPG CANFA
 ID LIPG CANFA STANDARD; PRT; 398 AA.
 AC P80035; O02857;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Triacylglycerol lipase, gastric precursor (EC 3.1.1.3) (Gastric
 lipase) (GL).
 GN LiPP.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Stomach;
 RC MEDLINE=99450174; PubMed=10520456;
 RA Vaganay S., Joliff G., Bertaux O., Toselli E., Devignes M.D.,
 RA Benicourt C.;

"The complete cDNA sequence encoding dog gastric lipase.";
[2]
DNA Seq. 8:257-262(1998).

RT RP SEQUENCE OF 20-59.
RA MEDLINE=92037652; PubMed=1935982;
RA Carriere F., Moreau H., Raphael V., Laugier R., Benicourt C.,
RA Junien J.-L., Verger R.;
RT "Purification and biochemical characterization of dog gastric
RT lipase.";
RL Eur. J. Biochem. 202:75-83(1991).
CC -1- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O = diacylglycerol + a
CC fatty acid anion.
CC -1- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

CC EMBL; Y13899; CAA74198.1; -;
DR PDB; 1K80; 20-MAR-02.
DR InterPro; IPR000073; A/b hydrolase.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Serestrs_site.
DR Pfam; PF04083; abhydro_lipase; 1.
DR Pfam; PF00561; abhydro_lipase; 1.
DR PROSITE; PS00120; LIPASE SER; 1.
KW Hydrolase; Lipid degradation; Glycoprotein; Signal; 3D-structure.
FT SIGNAL 1 19
FT CHAIN 20 398 TRIACYLGLYCEROL LIPASE, GASTRIC.
FT ACT_SITE 172 172 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 372 372 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CARBOHYD 34 34 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 39 39 I -> T (IN REF. 2).
SQ SEQUENCE 398 AA; 45130 MW; E04D62F751B3E386C CRC64;

Query Match 55.0%; Score 1141.5; DB 1; Length 398;
Best Local Similarity 56.5%; Pred. No. 9.5e-94;
Matches 210; Conservative 67; Mismatches 90; Indels 5; Gaps 3;

QY 4 PTKAVDPEAFNISEIIOHOGYPCPEEVEVATEDGYILSVNRIPRGLVQPKTGSRPVWLL 63
DB 26 PT---NEEVNIIISQITMGYPABEYEVVEDGYILGIDRIPYGRKNSENIGRRPVAF 82
QY 64 QHGLVGASNIWISNLPNNSLGLFADAGFDVWNGNSRGNWSRKHKTLSIDQDEFWAF 123
DB 83 QHGLASATNWIWISNLPNNSLGLFADAGFDVWNGNSRGNWTWARRNLYYSPDSVEFWAF 142
QY 124 DEMARFDPVAVINILORTGOEKIYVYGSQGTMTGFIAPSTWPELAOKIKYFALAP 183
DB 143 DEMAKYDLPATIDFILKKTGDKLHYVGHSGQGTIGFIAPSTWPKAKRIKTYALAP 202
QY 184 TVKHAISPGTKFLLPDMWIKGLFGKKEFLYQTRFLRQ-LVIVLCGGVILDDQICSNIM 242
DB 203 TVKYTETLLNKLMLVPSFLKIFGNKIF-YPHFPDQFLATEVCSRETVDLCSNALFI 261
QY 243 LGGTNTNMNSRASVTAHTLACTSVQNTILHWSQAVNSGELRAPFGWSTKMLEKNQ 302
DB 262 ICGPDTNLANSLRDVYLSNPAGTSVQNTILHWSQAVNSGELRAPFGWSTKMLEKNQ 321
QY 303 TPVRYRVDMPVPTAMWTGGDMLSNPDKVLLSVYTNLYHKNIPENAHVDFIWLDA 362
DB 322 MPPYNTLDMHVPVAVNGNDLLADPHVDLILLKLPNLIYHKKIPPYNELDFIWA 381
QY 363 PHRMNIEIHL 374
DB 382 POAVYNEIVSNM 393

RESULT 6

LIPG_BOVIN
ID LIPG_BOVIN STANDARD; PRT; 397 AA.
AC Q29458;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Triacylglycerol lipase, pregastric precursor (EC 3.1.1.3) (Pregastric
DE lipase) (GL) (Pregastric esterase) (PGE).
GN LIPF.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=torque serous gland;
RX MEDLINE=95011625; PubMed=7926811;
RA Timmermans M.Y.J., Kupers L.P., Teuchy H.;
RT "The cDNA sequence encoding bovine pregastric esterase.";
RL Gene 147:259-262(1994).
CC -1- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O = diacylglycerol + a
CC fatty acid anion.
CC -1- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

CC EMBL; L26319; AAA57037.1; -;
DR PIR; JC4017; JC4017.
DR InterPro; IPR000073; A/b hydrolase.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Serestrs_site.
DR Pfam; PF04083; abhydro_lipase; 1.
DR Pfam; PF00561; abhydro_lipase; 1.
DR PROSITE; PS00120; LIPASE SER; 1.
KW Hydrolase; Lipid degradation; Glycoprotein; Signal.
FT SIGNAL 1 19
FT CHAIN 20 397 TRIACYLGLYCEROL LIPASE, PREGASTRIC.
FT ACT_SITE 171 171 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 371 371 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CARBOHYD 33 33 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 270 270 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 326 326 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 397 AA; 45231 MW; F68977DED585EE36 CRC64;

Query Match 54.2%; Score 1125; DB 1; Length 397;
Best Local Similarity 55.8%; Pred. No. 2.8e-92;
Matches 208; Conservative 64; Mismatches 99; Indels 2; Gaps 2;

QY 7 AVDPFAFNIGSEIIOHOGYPCPEEVEVATEDGYILSVNRIPRGLVQPKTGSRPVWLLQHG 66
DB 25 AKVPEASNVNQMSLYWGYPSMHKVTADGYILQVYRIPHGKNNANHLGQRPVFLQHG 84
QY 67 LVGGASNIWISNLPNNSLGLFADAGFDVWNGNSRGNWSRKHKTLSIDQDEFWAFSD 126
DB 85 LIGSATNWIWISNLPNNSLGLFADAGFDVWNGNSRGNWTWARRNLYYSPDSVEFWAF 144
QY 127 ARFDLPVAVINILORTGOEKIYVYGSQGTMTGFIAPSTWPELAOKIKYFALAP 186
DB 145 AEDLPSTIDFILTARTGQKLYVGHSGQGTIGFIAPSTWPELAOKIKYFALAP 204
QY 187 HAKSPGTKFLLPDMWIKGLFGKKEFLYQTRFLRQ-LVIVLCGGVILDDQICSNIM 245
DB 205 YTKSLFNKALIPFLFKIIFGDKMF-YPHFTLQFLGVENCSRETLDVLCKNALFAITG 263

QY 246 FNTNNMNSRASVYAAHTLACTSVQNLHWSQAVNSGELRAPDWSGSETKNLEKCNQPTPV 305
 Db 264 VOKNFNMRLDVIYIAHNPAGTSVQNLHWRQAVSGKFPQADGAPYQNLHMYHQPTPP 323
 QY 306 RYVRDVTMTVMTWGTGGQDLNPNEDVOMLLSEVNTLIYHKNIPWVAHVDTWGLDAPHR 365
 Db 324 IYNTAMNVPFAVWSADNLLADPDQVDVFLSKSLNLIYHKEIPYNNHLDPIWAMDAPQE 383
 QY 366 MYNEIHLMOQEE 378
 Db 384 VYNEIVSLMAEDK 396

RESULT 7
 LIP3 HUMAN STANDARD; PRT; 398 AA.
 ID LIP3 HUMAN STANDARD; PRT; 398 AA.
 AC P07098;
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Triacylglycerol lipase, gastric precursor (EC 3.1.1.3) (Gastric
 lipase) (GL).
 GN LIPF.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP MEDLINE=97299724; PubMed=3304425;
 RA Bodmer M.W., Augal S., Yarranton G.T., Harris T.J.R., Lyons A.,
 King D.J., Pieroni G., Riviere C., Verger R., Lowe P.A.;
 RT "Molecular cloning of a human gastric lipase and expression of the
 enzyme in yeast."
 RL Biochim. Biophys. Acta 909:237-244 (1987).
 RN [2]
 RP SEQUENCE OF 20-45.
 RX MEDLINE=89325292; PubMed=2753032;
 RA Bernbeck S., Blaeckberg L.;
 RT "Human gastric lipase. The N-terminal tetrapeptide is essential for
 lipid binding and lipase activity."
 RL Eur. J. Biochem. 182:495-499 (1989).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
 RX MEDLINE=9287897; PubMed=10358049;
 RA Roussel A., Canaan S., Eglhoff M.P., Riviere M., Dupuis L., Verger R.,
 Cambillau C.;
 RT "Crystal structure of human gastric lipase and model of lysosomal
 acid lipase, two lipolytic enzymes of medical interest."
 RL J. Biol. Chem. 274:16995-17002 (1999).
 CC -1- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O = diacylglycerol + a
 fatty acid anion.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X05997; CAA29413.1; --
 DR EMBL; X05997; CAA29414.1; ALT_INIT.
 DR EMBL; A01046; CAA00125.1; --
 DR EMBL; A12714; CAA01053.1; --
 DR PIR; S07145; S07145.
 DR PDB; 1HLG; 15-MAR-00.
 DR Genew; HGNC:5622; LIPF.
 DR MIM; 601980; --
 DR GO; GO:0008289; F:lipid binding activity; NAS.

GO; GO:0004806; F:triacylglycerol lipase activity; TAS.
 GO; GO:0006641; P:triacylglycerol metabolism; NAS.
 DR InterPro; IPR000073; A/B_hydrolase.
 DR InterPro; IPR000734; Lipase.
 DR InterPro; IPR000379; Ser_estrs_ssite.
 DR Pfam; PF004083; abhydro_lipase; 1.
 DR Pfam; PF00561; abhydro_lipase; 1.
 DR PROSITE; PS00120; LIPASE_SER; 1.
 KW Hydrolase; Lipid degradation; Glycoprotein; Signal; Polymorphism;
 KW 3D-structure.
 FT SIGNAL 1 19
 FT CHAIN 20 398 TRIACYLGLYCEROL LIPASE, GASTRIC.
 FT ACT_SITE 172 172 CHARGE RELAY SYSTEM.
 FT ACT_SITE 343 343 CHARGE RELAY SYSTEM.
 FT ACT_SITE 372 372 CHARGE RELAY SYSTEM.
 FT DISULFID 246 255
 FT CARBOHYD 34 34 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 327 327 N-LINKED (GLCNAC. . .).
 FT VARIANT 161 161 T -> A (IN ABSNP:814628).
 FT /FTID=VAR 011947.
 SQ SEQUENCE 398 AA; 45237 MW; CD3EE1621C014F0F CRC64;
 Query Match 53.4%; Score 1109; DB 1; Length 398;
 Best Local Similarity 54.7%; Pred. No. 7.4e-31;
 Matches 202; Conservative 69; Mismatches 98; Indels 0; Gaps 0;
 QY 10 PEAFMNISEIIHQGVPCPEYEVATEDGYILSVNRIPIRGVLVQPKKTGSRPVLQHLVG 69
 Db 29 PEVTMNIISQITMYGVNEEYEVVTDGYILEVNRIPYKKNKSGNTGQRPVFLQGLLA 88
 QY 70 GASWISNLPNSLGFILADAGFDVWNSRGNAWKRKHKLISIDQDEFWAFSYDEWARF 129
 Db 89 SATWISNLPNSLAFILADAGYDVLGNSRGNTWARNLYSPDSVEFWAFSDEWAKY 148
 QY 130 DLPVINFILQKQEKIYVGVYSGQTMGTGFIASFPELAQIKWYFALAPIATVKHAK 189
 Db 149 DLPATIDFVKKTKQKQLHYVHSGQTTIGTGFIAFNPISLAKIKTFYALPAPVATVKYTK 208
 QY 190 SPGTFKLLPDMIKGLFGKEFLYQTRFLRQLVILCGQVILDOICSNMLLLGNTN 249
 Db 209 SLINKLRFVQSLFKFIFGDKIFVPHNFDFQFLATEVCSREMLNLLCSNALFICGFDK 268
 QY 250 NMNMSRASVYAAHTLACTSVQNLHWSQAVNSGELRAPDWSGSETKNLEKCNQPTPVRYV 309
 Db 269 NNTSRDLVLSHPAGTSVQNLHWRQAVSGKFPQADGAPYQNLHMYHQPTPPYV 328
 QY 310 RDMVTPTAMTGGQDLNPNEDVOMLLSEVNTLIYHKNIPWVAHVDTWGLDAPHRMYNE 369
 Db 329 TAMNVPFAVWNGKDLADPDQVDVFLSKSLNLIYHKEIPYNNHLDPIWAMDAPQEVYND 388
 QY 370 IHLMOQEE 378
 Db 389 IVSMISEDK 397

RESULT 8
 LIP3 DROME STANDARD; PRT; 394 AA.
 ID LIP3 DROME STANDARD; PRT; 394 AA.
 AC Q46108;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lipase 3 precursor (EC 3.1.1.-) (DmLip3).
 GN LIP3 OR CG8823.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.

STRAIN-Canton-S;
MEDLINE=98227315; PubMed=9566193;
Piatillo D., Manzi A., Tino A., Pilo Boyl P., Graziani F., Malva C.;
"The Drosophila melanogaster lipase homologs: a gene family with
tissue and developmental specific expression.";
J. Mol. Biol. 276:877-885(1998).
[2]
SEQUENCE FROM N.A.
STRAIN=Berkely;
MEDLINE=20196006; PubMed=10731132;
Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champe M., Pfeiffer B.D.,
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
Abril J.F., Agbayani A., An H.-J., Andrews-Fianknoch C., Baldwin D.,
Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
Beeson K.Y., Benos P.V., Berland B.P., Bhandari D., Bolshakov S.,
Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
Burtis K.C., Busam D.A., Butler K., Cadiou E., Center A., Chandra I.,
Cherry J.N., Cawley S., Dahlke C., Davenport L.B., Davies P.,
de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
Fowler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
Jatani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
Lasko P., Lei Y., Levitsky A.C., Li J., Li Z., Liang Y., Lin X.,
Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
Reinart K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
Shue B.C., Siden-Klamis I., Simpson M., Skupski M.P., Smith T.,
Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
"The genome sequence of *Drosophila melanogaster*.";
Science 287:2185-2195(2000).
-!- TISSUE SPECIFICITY: FAT BODY.
-!- DEVELOPMENTAL STAGE: ONLY AT LARVAL STAGES.
-!- SIMILARITY: PARTIAL WITH OTHER LIPASES (PANCREATIC, GASTRIC,
HEPATIC, LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.). ALSO SIMILAR TO
LEPIDOPTERAN EGG-SPECIFIC AND YOLK PROTEINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
or send an email to license@isb-sib.ch).

EMBL; Y14367; CAA74737.1; -;
EMBL; AE003699; AAF54935.1; -;
Flybase; FBGN0023495; Lip3.
InterPro; IPR000073; A/b_hydrolase.
InterPro; IPR000734; Lipase.
InterPro; IPR000379; Ser_estra_site.
Pfam; PF04083; abhydro_lipase; 1.
Pfam; PF00561; abhydro_lase; 1.
PROSITE; PS00120; LIPASE_SER; 1.
HydroLase; Lipid degradation; Signal; Glycoprotein.
FT
SIGNAL 1 20
POTENTIAL.

FT	CHAIN	21	394	LIPASE 3.	
FT	ACT_SITE	164	164	CHARGE RELAY SYSTEM (BY SIMILARITY).	
FT	ACT_SITE	369	369	CHARGE RELAY SYSTEM (BY SIMILARITY).	
FT	CARBOHYD	131	131	N-LINKED (GLCNAC..) (POTENTIAL).	
SQ	SEQUENCE	394 AA;	44901 MW;	A718D1D743673802 CRC64;	
	Query Match	31.9%;	Score 663;	DB 1;	Length 394;
	Best Local Similarity	37.0%;	Pred. No. 3e-51;		
	Matches 137;	Conservative	80;	Mismatches 139;	Indels 14; Gaps 7;
QY	18	EITQHOGQCEBEYEVATEDG	YILSVNRIPRGLVQPK--	KTGSRPVVLLHGLVGGASNWI	75
DB	30	ERIEDDGYPMERHEVVTSD	NYILTMHRIP--	YSPKGTGESSNRPVAFMHGMLSSSSDWW	86
QY	76	SNLPNNLSGLFILADAFD	VMGNSRGNVSRKUTLSID	ODEFWATSYDEMARFDLP	PAVI 135
DB	87	LMGPFERSLAYMLADAGY	DDVMGNARGNTYSKAHKY	PTWQIIPNPFNSWNEIGMYD	VPAMI 146
QY	136	NFILOKTGQEKIYVGVSG	QTMTGFTAFSTMPLEAOK	IKMYFALAPATATVKHAKS	PGTK- 199
DB	147	DYVLAKTGGQOVQVGHSG	QTTVYLVVMSRPEYNDKIS	AHLLGPLAAYVGNWKSPL	TRA 206
QY	195	---FLLLPDMIKGLFGK	EFLYQTRPLRQLVYL	CGQVI-LDOI	CSNIMLLGGFTWNN 250
DB	207	FAPILGPNAIVE-VCGSM	EFMPENKFKQDLGIEMCO	ATSPYADMCANEIFLIGGY	DTQE 265
QY	251	MNMSRASVVAATHLAGT	SVQNILHWSQAVNSGLR	AFDWMGSEFKLEKCNQPT	PVRVR 310
DB	266	LDYELLEHIKATSPAG	ASVNVQNLIHFQEYNSG	KRFEDY-TALRN	PYEYSYPPDPYKX 324
QY	311	DMTVPTAMTGGODLNS	PNPDVKKLLSEVTNLIY	HKNP--EWAHVD	DFIWGLDAPHRMYN 368
DB	325	NAKAPVLLYYGANDW	CMCDVSDVRKLRDEL	PNMALDYLVPPEKWAHL	DFIWGTTEARKYVD 384
QY	369	EIHLHMQQEE 378			
DB	385	EVLKMQQSYE 394			
RESULT 9					
ID	LIP1_DROME	STANDARD;	PRF;	439 AA.	
AC	O46107; Q9VXR6;				
DT	16-OCT-2001 (Rel. 40, Created)				
DT	16-OCT-2001 (Rel. 40, Last sequence update)				
DT	15-SEP-2003 (Rel. 42, Last annotation update)				
DE	Lipase 1 precursor (EC 3.1.1.-) (DmLip1).				
GN	LIP1 OR CG7279.				
OS	Drosophila melanogaster (Fruit fly).				
OC	Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;				
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
OC	Ephydroidea; Drosophilidae; Drosophila.				
OX	NCBI_TaxID=7227;				
RN	[1]				
RP	SEQUENCE FROM N.A., FUNCTION, TISSUE SPECIFICITY, AND DEVELOPMENTAL				
RP	STAGE.				
RC	STRAIN=Canton-S;				
RC	MEDLINE=98227315; PubMed=9566193;				
RA	Pistillo D., Manzi A., Tino A.,				
RA	Adams M.D., Celniker S.E., Li P.W., Hoskins R.A., Galle R.F.,				
RA	Ananides P.G., Scherer S.E.,				
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,				
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,				
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,				
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,				
RA	Abriel J.F., Achavani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,				
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,				
RA	Adams M.D., Celniker S.E., Li P.W., Hoskins R.A., Galle R.F.,				
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,				
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,				
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,				
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,				
RA	Abriel J.F., Achavani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,				
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,				
RA	Adams M.D., Celniker S.E., Li P.W., Hoskins R.A., Galle R.F.,				
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,				
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,				
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,				
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,				
RA	Abriel J.F., Achavani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,				
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,				
RA	Adams M.D., Celniker S.E., Li P.W., Hoskins R.A., Galle R.F.,				
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,				
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,				
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,				
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,				
RA	Abriel J.F., Achavani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,				
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,				
RA	Adams M.D., Celniker S.E., Li P.W., Hoskins R.A., Galle R.F.,				
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,				
RA	Sutton G.G., Wortman J.R., Yandell M.D.,				

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Bueam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos J.M., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K.J., Doup L.E., Downes M., Duggan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
RA Fesler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laško P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas P., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RA "The genome sequence of *Drosophila melanogaster*,"
RL Science 287:2185-2195 (2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley; TISSUE=Embryo;
RX MEDLINE=22426066; PubMed=12537569;
RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA George R.A., Guarin H., Kronmiller B., Pacleb J.M., Park S., Wan K.H.,
RA Rubin G.M., Gelniker S.E.;
RT "A *Drosophila* full-length cDNA resource,";
RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8 (2002).
CC -!- FUNCTION: COULD BE A DIGESTIVE ENZYME.
CC -!- SUBCELLULAR LOCATION: Secreted (Probable).
CC -!- TISSUE SPECIFICITY: In 14 hour embryos expression is seen in the
CC foregut/midgut boundary.
CC -!- DEVELOPMENTAL STAGE: Expressed from 14 hour embryos through to
CC adulthood. There is a weak maternal contribution to early embryos.
CC -!- SIMILARITY: PARTIAL WITH OTHER LIPASES (PANCREATIC, GASTRIC,
CC HEPATIC, LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.). ALSO SIMILAR TO
CC LEPIDOPTERAN EGG-SPECIFIC AND YOLK PROTEINS.
CC -!- CAUTION: It is uncertain whether Met-1 or Met-7 is the initiator.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y14366; CAAY74736.1; ALT_INIT.
DR EMBL; AB003629; AAF52994.1; -.
DR EMBL; AY075506; AAL68315.1; -.
DR FlyBase; FBgn0023496; Lip1.
DR InterPro; IPR000073; A/b hydrolase.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser esters site.
DR Pfam; PF04083; anhydrolipase; 1.
DR Pfam; PF00561; anhydrolipase; 1.
DR PROSITE; PS00120; LIPASE SER; FALSE NEG.
KW Hydrolase; Lipid degradation; Signal; Glycoprotein.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 439 LIPASE 1.
FT DOMAIN 30 44 POLY-GLU.

FT ACT SITE 197 197 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT SITE 393 393 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CARBOHYD 124 124 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 151 151 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 346 346 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 379 379 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 426 426 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 10 10 L -> I (IN REF. 1).
FT CONFLICT 213 213 Y -> F (IN REF. 1).
FT CONFLICT 412 412 Q -> E (IN REF. 1).
SQ SEQUENCE 439 AA; 50660 MW; 9E32B20BEAE93E3F CRC64;
Query Match 26.3%; Score 546; DB 1; Length 439;
Best Local Similarity 31.8%; Pred. No. 8.6e-41;
Matches 123; Conservative 83; Mismatches 157; Indels 24; Gaps 7;
QY 6 KAVDPEAFNISEIIHQGYCFEVEVATEDGYILSVNRIPLGLVQPKKTSRPPVILLOH 65
DB 58 KNIKODSTLSVDKLIATKYGESEVHVTTEDGYILTHRI-----RKQA-PFFLLOH 109
QY 66 GLVGASGWNISNLNNSLGLADAGFDVMNGSRGNMWSRKHKTLISIDQDEFWAFSYDE 125
DB 110 GLVDSSAGFVGMGNVSLAYLLADHNYDVMLGNARGNRYSRNHTTLDPDESKFDFSWHE 169
QY 126 MARFPLPAVINFILOKTOEKLIVYGYSGTGMFIAPSTPELAQKIKMYFALAPIATV 185
DB 170 IGMVDLPAMIDHVLKVTGFFKLVHAGHSQGTCTFFVCMSEMPAYNDKVVSMQALAPAVYA 229
QY 186 KHAKS-PGTFKLLPDMMIKGLFGKKEFLYQTRFLQRLVILYLCQGVILDCISIMILLG 244
DB 230 KETEDHPYIRALSIFYNSLVGSSIREMNGEFLCRMT-----EETERLCIEAVFGIV 283
QY 245 GNTNNMMSRASVYAHTLAGTSVQNLHWSQVNSGELAPFDGSETKLEKCNQPTP 304
DB 284 GNNWNEFRKMPFVILGHYPAGVAAKQYHFIQIKSRFPAPYSY-SNNKMQLYRDHLP 342
QY 305 VYRYVRDMVTPTAMWTGGQDLNSNPEDVKMLLSEVTLNLIYHKNP--EWAHVDFIWGLDA 362
DB 343 PRYNLSLVTVPFVYVYSTNDLLCHPQVDSMCDLGNVTGKLYVPOKEFNEMDFLWIDV 402
QY 363 PHRMVNEIHLN-----QQEETNLQ 383
DB 403 RKMLYRMLQVLGKVPFGSPSEANRSR 429
RESULT 10
TG1_YEAST STANDARD; PRT; 548 AA.
AC P34163;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Triglyceride lipase-cholesterol esterase (EC 3.1.1.-).
TG1 OR YKLI40W OR YKLI5.
OS Saccharomyces cerevisiae (Baker's yeast).
CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
CX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92245761; PubMed=1574929;
RA Abraham P.R., Mulder A., Van't Riet J., Planta R.J., Raue H.A.;
RT "Molecular cloning and physical analysis of an 8.2 kb segment of
RT chromosome XI of *Saccharomyces cerevisiae* reveals five tightly linked
RT genes,"
RL Yeast 8:227-238 (1992).
RN [2]
RP SEQUENCE FROM N.A.
RA Rad M.R., Xu G., Kirchner L., Fritz C., Keuchel H., Hollenberg C.P.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBSJ databases.
CC -!- SIMILARITY: Belongs to the AB hydrolase superfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; 225464; CRA80958.1; -
 DR EMBL; 228140; CRA81981.1; -
 DR PIR; S37969; S37969.
 DR SGD; S0001623; TGL1.
 DR GO; GO:000581; C:lipid particle; IDA.
 DR InterPro; IPR000073; A/B_hydrolase.
 DR InterPro; IPR000379; Ser_eastr site.
 DR Pfam; PF04083; abhydro_lipase; 1.
 DR Pfam; PF00561; abhydro_lipase; 1.
 DR KW Hydrolase; Lipid degradation.
 SQ SEQUENCE 548 AA; 62979 MW; 32D1F230701CB083 CRC64;

Query Match 20.4%; Score 422.5; DB 1; Length 548;
 Best Local Similarity 31.6%; Pred. No. 1.1e-29;
 Matches 119; Conservative 66; Mismatches 127; Indels 65; Gaps 15;
 QY 28 BEYEVEDGVLISVNRIPRGVOP--KKTGSRPVVLLQHLGVLGASNWI-SNLP--NNSLG 84
 DB EDHLVREEDNVLILHRIP-----PISKRRFNKVVVILHGLLMCSVWCNIRHKNLP 132
 QY 85 FILDAGFDVVMGNSRGNARSKHTLSIDODEFAPSYDPMARFDLPVAVINFILOKTGQ 144
 DB FVLHDLGVVVMGNNRGNKYSTAHLNPKPKSKWDFSIDFAFPDIPNSIERFIDITKV 192
 QY 145 EKIVYVGYSGCTWGTAFSTWPLAKIKMYFALAPIATVK--HAK-----SPGTK 194
 DB DKVICIGFSQSGAQMFAAFSLSEKLNKRVSHFIAIAPMTKGLHNRIVDTLAKSPGFM 252
 QY 195 FLLPLDMIMKGLFGKKEPLYOT----RFLRLVLYLCQVILDOICSNIMLLGFGFTNN 250
 DB 253 YL-----PFGRKIVLPSAVIQRWTLHPTLENLC-----ID-IANKIL-----FNWKS 293
 QY 251 MMW-----SRASVYAAHTLAGTSVQNLHWSQAVNSGELRAPD-----WGSETKNLEK 298
 DB 294 FNILPRQKIASYAKLYST-----TSVKSIVHWFQILRSQKFPQFESDNMLSLTRPYQI 348
 QY 299 CNQPTPVRYVRVDMTPTAMTGTQDMLSNPDEVDKMLLSEVTLNLIYHKNTPWAHVDVFIW 358
 DB 349 ANFPRT-----NIKIPILLYGGISDLSVDVDMKNLP--FNSVFDVKVDNYEHLDLIW 401
 QY 359 GLDAPHRMNIEIHLMQ 375
 DB 402 GKADATLVIAKVLRFIE 418

RESULT 11
 YD94 METJA
 ID YD94 METJA STANDARD; PRT; 987 AA.
 AC Q58789;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein MJ1394.
 GN MJ1394.
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 OC Methanocaldococcaceae; Methanocaldococcus.
 OX NCBI_TaxID=2190;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=8688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,

RA Scott J.L., Geoghegan N.S.M., Weigman J.P., Fuhrmann J.L., Nguyen D.,
 RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 jannaschii.";
 RL Science 273:1058-1073 (1996).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: TO M.JANNASCHII MJ1393 AND A.FULGIDUS AP2028.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; U67579; AAB99404.1; -
 DR EMBL; A64474; A64474.
 DR TIGR; MJ1394; -
 KW Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 12 32 POTENTIAL.
 FT TRANSMEM 958 978 POTENTIAL.
 SQ SEQUENCE 987 AA; 112360 MW; D1E628FBE28CA86D CRC64;

Query Match 5.4%; Score 111.5; DB 1; Length 987;
 Best Local Similarity 24.1%; Pred. No. 0.1;
 Matches 93; Conservative 46; Mismatches 124; Indels 123; Gaps 27;
 QY 17 SRIHQHGYPCCEEVEVATEDGYILSVNRIPRG--LVQPKTKGSRPVVLLQHLGVLGASN 73
 DB 408 SSIQSHSGNYSLEKYGIST-----SLANDPNGYKLPKEIG-RDVVI--SGWYRPSN 458
 QY 74 WISNLPNNSLGFILADAGFDVVMGNSRGNARSKHT--LSIDQ-----DEFWA 120
 DB 459 W-GGGPIDRIG--LEDENFD-----GVSEFVNHYNSVSDIRATNGNPTSEISPVYWN 508
 QY 121 PSYDEMARFDLPVAVINFILOKTGQEKIYVYGSGTTFGFIAPSTMPPLAKIKMYFALA 180
 DB 509 PPEDEWYFEL-----KI-----YSNGT-----ITFST-----YYONG 536
 QY 181 PT-ATVXGIAKSPGTFK-----LLLPDMNMGK-----LFGKKEFLYQTRFLRLQVLY 225
 DB 537 SLAATVSTIDNTYTKFDRVHVGIVGVYVVDLEVNSKNFDYGDKNWKL-----586
 QY 226 LCGQVILDOICSNIMLLGG-----FNTNNMNSRASVYAAHTL-----AGTSVQNIL- 273
 DB 587 ---EITSANSSSEGTAVLFDGDFYKDYNTSNLN---AINWNTNITLNSNDSATLVFNYLG 640
 QY 274 HWSQAVNSGELRAFWDGSETKNLEKCNQPTFVR--YRVRDMTPTAMWTGQ--DWLSN- 328
 DB 641 NYISYERDNLAKYGFAPKILFNYNGTNTNTSIKGYASGSYSISTDGHGTGGEINIWIENV 700
 QY 329 --PEDVKMLLSEVTLNLIYHKNIPewa 352
 DB 701 TPKNDAKSYSFNLITNL-----NI--WA 720

RESULT 12
 ORC2 CAEEL STANDARD; PRT; 430 AA.
 ID ORC2 CAEEL
 AC Q21037;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Origin recognition complex subunit 2 (C6orc2).
 GN ORC-2 OR F59E10.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=96099401; PubMed=7502077;
RA Gavin K.A., Hidaka M., Stillman B.;
RT "Conserved initiator proteins in eukaryotes."; Science 270:1667-1671(1995).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=Bristol N2;
RA Swinburne J.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RN REVISIONS.
RP Jones S.J.M.;
RA Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RL -1- FUNCTION: COMPONENT OF THE ORIGIN RECOGNITION COMPLEX (ORC) THAT BINDS ORIGINS OF REPLICATION. IT HAS A ROLE IN BOTH CHROMOSOMAL REPLICATION AND MATING TYPE TRANSCRIPTIONAL SILENCING. BINDS TO THE ARS CONSENSUS SEQUENCE (ACS) OF ORIGINS OF REPLICATION IN AN ATP-DEPENDENT MANNER.
CC -1- SUBUNIT: ORC IS COMPOSED OF SIX SUBUNITS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: BELONGS TO THE ORC2 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; U40270; AAC46954.1; -;
DR EMBL; Z36949; CAAB5415.1; -;
DR PIR; T23001;
DR WormPep; F59E10.1; CE11492.
DR Pfam; PF04084; ORC2; 1.
KW DNA replication; Nuclear protein.
SQ SEQUENCE 430 AA; 49319 MW; 54FC086B4AD9670 CRC64;
Query Match 5.1%; Score 105; DB 1; Length 430;
Best Local Similarity 21.9%; Pred. No. 0.12;
Matches 61; Conservative 42; Mismatches 87; Indels 88; Gaps 13;
QY 52 PKTGRSPVVLQGLVGGASNWSLNNPNN-----SLGF-----TLADAGFDVM--MGN 98
DB 25 PEKEGR-----QKKTNGKNAASRLQSNLEBDEQLGFEDETVSMASAIENYFMQK 78
QY 99 S-----RGNAWRKH-----TLSDQDEFWAFSYDEMARFDPVAVINILQKTOGKIY 148
DB 79 SASRMNNAKRRGRAGNNTBEIDDEISNAITDPTKCDLPGRNITKDNTEPEK 138
QY 149 YVGYSQGTWTFIAFTMPPELAQIKWYFALAPIATVYHAKSPGKFLLLPDMWINGLFG 208
DB 139 RLEHLADNDFG-----KWKLYLAAG-----FNILLHGVGS 168
QY 209 KKEFLYOTRPLQVILVLCQVILDCQSNIMLLGGFTN---NNMSEASVYAHTLA 265
DB 169 KRDLV--TERENELSDTYRVRDARKGLNVRVLLGAINENMKLNCNVRG-----217
QY 266 GTSYQNTLHWSQV---NSGEL-----RAPDWGSE 292
DB 218 ---QSTISWARSIRRMNSQQLIILIIDNEIAPDWRS 251

RESULT 13
PAPA CAVPO
ID PAPA CAVPO STANDARD; PRT; 436 AA.
AC P70683;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Platelet-activating factor acetylhydrolase precursor (BC 3.1.1.47)
DE (PAF acetylhydrolase) (PAF 2-acylhydrolase) (LDL-associated

phospholipase A2) (LDL-PLA(2)) (2-acetyl-1-alkylglycerophosphocholine esterase) (1-alkyl-2-acetyl-glycerophosphocholine esterase).
GN PLA2G7 OR PAFAH.
OS Cavia porcellus (Guinea pig).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
CN NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RP STRAIN=Hartley; TISSUE=Liver;
RX MEDLINE=97103479; PubMed=8947850;
RA Karasawa K., Kuge O., Kawasaki K., Nishijima M., Nakano Y., Tomita M., Yokoyama K., Setaka M., Nojima S.;
RT "Cloning, expression and characterization of plasma platelet-activating factor-acetylhydrolase from guinea pig."; J. Biochem. 120:1838-844 (1996).
RL -1- FUNCTION: MODULATES THE ACTION OF PLATELET-ACTIVATING FACTOR (PAF) BY HYDROLYZING THE SN-2 ESTER BOND TO YIELD THE BIOLOGICALLY INACTIVE LYSO-PAF. HAS A SPECIFICITY FOR SUBSTRATES WITH A SHORT RESIDUE AT THE SN-2 POSITION. IT IS INACTIVE AGAINST LONG-CHAIN PHOSPHOLIPIDS.
CC -1- CATALYTIC ACTIVITY: 2-acetyl-1-alkyl-sn-glycero-3-phosphocholine + H(2)O = 1-alkyl-sn-glycero-3-phosphocholine + acetate.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: Plasma.
CC -1- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; D67037; BAA11054.1; -;
DR PIR; JC5021; JC5021.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR005065; PAF-AH_P.II.
DR InterPro; IPR000379; Ser estersite.
DR Pfam; PF03403; PAF-AH_P.II.1.
DR PROSITE; PS00120; LIPASE_SER; 1.
KW Hydrolase; Lipid degradation; Glycoprotein; Signal.
FT SIGNAL 1 21
FT CHAIN 22 436
FT ACT_SITE 271 271
FT ACT_SITE 294 294
FT ACT_SITE 349 349
FT CARBOHYD 76 76
FT CARBOHYD 200 200
FT CARBOHYD 324 324
SQ SEQUENCE 436 AA; 49062 MW; C35D96E392FE11 CRC64;
Query Match 4.7%; Score 98.5; DB 1; Length 436;
Best Local Similarity 20.5%; Pred. No. 0.48;
Matches 71; Conservative 49; Mismatches 128; Indels 99; Gaps 16;
QY 44 RIPGLVQPKTKGR-PVLLQHLQGVGASNWSLNNPNSLGLADAGPDV-----94
DB 129 KVPKWSPLKTKGKPLIIFSHGL--GAFRSI-----YSAIGIELASHGIVAAVEHRDE 182
QY 95 -----WMGNSRGNASRKHKTLISIDODE-----FWAFSYDEMAR 128
DB 183 SAAATYTFQAPAAESGNRSWIYKGNLETERKQLRQRCGECQALSLLSIDEGEP 242
QY 129 FDLPAVINFLQK-----TGQEKIYVYGYSQGTWTFIAFTMPPELAQIKWYFAL-----A 180
DB 243 VKVLDLNFDDIQKGLSDRSKVAIGHSGF--GATVIQTLSE-DQFRFCGIALDPWMF 298
QY 181 PIATVYHAKSPGKFLLLPDMWINGLFGKK-EFLYQTRFLRLQV-----IYL 226
DB 299 FVGSDVHSHKIPQLPFFINSEYFOSANDTKKIEFYQPKERKMIKVGSHVHNFVDTFA 358

ACT SITE 258 258 BY SIMILARITY.

FT STRAND 3 9

FT TURN 10 11

FT STRAND 12 22

FT STRAND 26 30

FT TURN 33 34

FT STRAND 37 40

FT HELIX 41 49

FT TURN 50 51

FT STRAND 53 57

FT TURN 60 61

FT TURN 63 64

FT HELIX 74 88

FT TURN 89 89

FT STRAND 93 98

FT TURN 99 99

FT HELIX 100 112

FT TURN 115 116

FT STRAND 117 123

FT STRAND 131 131

FT TURN 137 138

FT STRAND 141 141

FT HELIX 143 155

FT HELIX 157 168

FT TURN 169 170

FT HELIX 171 174

FT STRAND 175 175

FT TURN 176 178

FT STRAND 179 179

FT HELIX 181 193

FT HELIX 196 205

FT TURN 206 207

FT TURN 211 213

FT HELIX 214 216

FT STRAND 221 226

FT TURN 227 228

FT HELIX 234 236

FT TURN 237 237

FT HELIX 238 244

FT TURN 246 247

FT STRAND 249 253

FT TURN 254 255

FT TURN 258 259

FT HELIX 260 263

FT TURN 264 264

FT HELIX 265 278

SQ SEQUENCE 278 AA; 30354 MW; 8C97A87251FBEDES CRC64;

Query Match 4.5%; Score 93.5; DB 1; Length 278;

Best Local Similarity 22.6%; Pred. No. 0.72;

Matches 72; Conservative 36; Mismatches 114; Indels 97; Gaps 17;

QY 58 RPVLLQHLVGGASNMISLPNNLSGLFADAGFDVWNGNRGNAMSRKHKTLISIDQDE 117

Db 25 QPVLLI-HGFFLSGHSW-----ERQSAALLDAGYRIVTYDRRGFGQSQPTT----- 70

QY 118 FWAFSYDEMARFDPAPVINFILQKTQEKIYVGYSGQTTMGFTAFSTMPELAKIKMYF 177

Db 71 --GYDYDTFA-----ADLNTVLETLDLQDAVLGVFSMGT-----GEVARYVSSY- 112

QY 178 ALAPIATYKHAKSPGPKELLAPDMIKLPGKKEPLVQTRFLRQLVIVLCGVILQICS 237

Db 113 GTARIKAVAFASL-EPFLKTKTDNPDGA-APKEF-----FDGIVA 151

QY 238 NIMLLGGFNTNMMMSRASVYAHTLAGT-----SVQNIHWSQAVNSGELRA-----FD 288

Db 152 AVKADRYAFYTGFFN-----DFYNLDENLGTRISEEAVEN--SWNTAASGGFPFAAAAPT 205

QY 289 WGSEYK-NLEKNQPTFVRYVRDVTFTAMWTGGQDWLSNPEDVKMLLSEVTLNLIYKN 347

Db 206 WYTFDPRADIPRIDVPALIHGTGDRTLPI-----ENTARVFHKA 244

QY 348 IP--EWAHVDFIWLGLDAPH 364

Db 245 LPSAEYVEVE-----GAPH 258

Search completed: February 19, 2004, 14:21:13

Job time : 11:1144 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2004, 14:14:34 ; Search time 34.8598 Seconds
(without alignments)
1926.040 Million cell updates/sec

Title: US-10-042-431-47

Perfect score: 2247

Sequence: 1 MLETLROWIVSHREWMML.....IHLMQBETNLQGRCEAVL 423

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 19Jun03:*

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*

2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*

3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*

4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*

5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*

6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*

7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*

8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*

9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*

10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*

11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*

12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*

13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*

14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*

15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*

16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*

17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*

18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*

19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*

20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*

21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*

22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2247	100.0	423	22 AAB66065	Human TANGO 294.
2	2247	100.0	423	22 AABU08369	Amino acid sequenc
3	2174	96.8	409	22 AAE11931	Human CG162 (Or C5
4	2076	92.4	390	22 AAB66067	Human TANGO 294 ma
5	2013.5	89.6	423	23 AABG31894	Human lipid-associ
6	1991	88.6	398	23 AABG71000	Human lipase prote
7	1991	88.6	398	23 AAU99164	Human lysosomal ac
8	1536.5	68.4	371	23 AAU99165	Human lysosomal ac
9	1289	57.4	399	22 AAB90783	Human shear stress

10	1289	57.4	399	22 AAB66061	Human lysosomal ac
11	1282	57.1	392	23 AAO18227	Human lysosomal ac
12	1191.5	53.0	449	23 AAU30498	Novel human secret
13	1166	51.9	398	14 AAR37302	RGL precursor. Or
14	1165	51.8	221	22 AAB66068	Human TANGO 294 ex
15	1156.5	51.5	395	6 AAP50322	Rat lingual lipase
16	1143	50.9	395	23 AAE14744	Human triacylglyce
17	1143	50.9	399	23 AAU77496	Human lipid metabo
18	1141.5	50.8	379	15 AAR56870	Canine gastric lip
19	1141.5	50.8	379	17 AAW09382	Dog gastric lipase
20	1141.5	50.8	380	15 AAR56871	Canine gastric lip
21	1139	50.7	399	22 AAG67513	Amino acid sequenc
22	1136	50.6	397	23 AAE17308	Human lysosomal ac
23	1127	50.2	378	23 ABB76189	Kid goat pregastric
24	1119	49.8	398	7 AAP60724	Sequence of pregas
25	1119	49.8	398	7 AAP60658	Sequence of human
26	1119	49.8	398	17 AAW09383	Human gastric lipa
27	1119	49.8	398	22 AAB66086	Human lipase prote
28	1119	49.8	398	24 AAB96698	Human gastric lipa
29	1101	49.0	403	23 AAU77493	Human lipid metabo
30	1090	48.5	427	23 AAU98539	Human lysosomal ac
31	1034	46.0	401	23 AAG94641	Human NOV6b protei
32	1031	45.9	395	23 AAO18226	Human lysosomal ac
33	1031	45.9	395	24 AAG73144	Human lipase prote
34	1006	44.8	373	23 ABE53569	Human NOV1 protein
35	995.5	44.3	362	24 ABE96697	Human triacylglyce
36	977	43.5	365	23 AAE17307	Human lysosomal ac
37	909	40.5	331	24 ABR14626	Human DITHP secret
38	907	40.4	731	23 AAE14746	Human triacylglyce
39	901.5	40.1	390	23 AAG94640	Human NOV6a protei
40	783	34.8	144	22 AAB66070	Human TANGO 294 CY
41	781	34.8	276	23 AAE14745	Human triacylglyce
42	777	34.6	289	23 AAU77494	Human lipid metabo
43	745.5	33.2	280	23 AAE14743	Human triacylglyce
44	724.5	32.2	656	22 ABE26839	Novel human diagno
45	691	30.8	233	22 AAB61608	Human protein HP03

ALIGNMENTS

RESULT 1

AAB66065

ID AAB66065 standard; Protein; 423 AA.

AC AAB66065;

XX

XX 30-MAR-2001 (first entry)

DT

DE Human TANGO 294.

XX TANGO protein; INTERCEPT protein; neurological disorder;

KW central nervous system; focal brain disorder; bipolar affective disorder;

KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;

KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;

KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;

KW neuropsychiatric; psychoactive substance use; anxiety.

XX

OS Homo sapiens.

XX

XX WO200077239-A2.

PN

XX

PD 21-DEC-2000.

XX

XX 24-MAY-2000; 2000WO-US14858.

XX

XX 14-JUN-1999; 99US-0333159.

PR

XX (MILL-) MILLENNIUM PHARM INC.

PA

XX McCarthy SA, Fraser CC, Sharp JD, Barnes TM;

PI

XX WPI; 2001-032313/04.

DR

DR N-PSDB; AAF45131, AAF45132.

XX TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for

PT screening assays and diagnostic assays and for the treatment of

PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's

PT disease -

XX Claim 8; Fig 6; 359pp; English.

XX The present invention relates to TANGO or INTERCEPT proteins and coding

CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAF6031-B66057,

CC AAF6064-B66083 and AAF66085). The TANGO/INTERCEPT proteins and coding

CC sequences are useful for the treatment of neurological disorders such as

CC central nervous system (CNS) disorders, CNS-related disorders, focal

CC brain disorders, global-diffuse cerebral disorders and other

CC neurological and cerebrovascular disorders. The CNS disorders include

CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic

CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,

CC autonomic function disorders such as hypertension and sleep disorders,

CC neuropsychiatric disorders, psychoactive substance use disorders,

CC anxiety, and bipolar affective disorder.

XX SQ Sequence 423 AA;

Query Match 100.0%; Score 2247; DB 22; Length 423;

Best Local Similarity 100.0%; Pred. No. 6.8e-220;

Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLETLRQWIVSHRMWMLLILVAYMFQNRVNSVHMTKAVDPEAFNISEIHHQGYPC 60

Db 1 MLETLRQWIVSHRMWMLLILVAYMFQNRVNSVHMTKAVDPEAFNISEIHHQGYPC 60

Qy 61 EYEYVATEDGVILSVNRIIPRLGVQPKTGRPVVLLQHLVGGASNWSINLPNSLGFIL 120

Db 61 EYEYVATEDGVILSVNRIIPRLGVQPKTGRPVVLLQHLVGGASNWSINLPNSLGFIL 120

Qy 121 ADAGFDVWNGSRGNARSKHTLSIDQDFWAFSYDEMARFIDLPAVINEFILOKTQEKI 180

Db 121 ADAGFDVWNGSRGNARSKHTLSIDQDFWAFSYDEMARFIDLPAVINEFILOKTQEKI 180

Qy 181 YVGYSGQTGTFATSTMPBLAQIKMYFALAPIATVVKAKSPGTKFLLLPDMTKGLF 240

Db 181 YVGYSGQTGTFATSTMPBLAQIKMYFALAPIATVVKAKSPGTKFLLLPDMTKGLF 240

Qy 241 GKKEFLYQTRFLRQVLYLQGVILDOICSNIMLLGENTNNMNSRASVYAAHTLAGT 300

Db 241 GKKEFLYQTRFLRQVLYLQGVILDOICSNIMLLGENTNNMNSRASVYAAHTLAGT 300

Qy 301 SVQNILHWSQAVNSGELRAFQWGSSEKTNLEKCNQPTPVYRVDRMTVPTAMWTGGQDWLS 360

Db 301 SVQNILHWSQAVNSGELRAFQWGSSEKTNLEKCNQPTPVYRVDRMTVPTAMWTGGQDWLS 360

Qy 361 NPEDVKMLLSEVTNLIYHKNIPFAVHDFIWLGDAPHRMYNEIHLMQBETNLSQGRCE 420

Db 361 NPEDVKMLLSEVTNLIYHKNIPFAVHDFIWLGDAPHRMYNEIHLMQBETNLSQGRCE 420

Qy 421 AVL 423

Db 421 AVL 423

RESULT 2

ABU08369

ID ABU08369 standard; Protein; 423 AA.

XX AC ABU08369;

XX DT 03-JUN-2003 (first entry)

XX DE Amino acid sequence for human TANGO 294.

XX Human; TANGO 294; INTERCEPT; cellular process; tissue typing;

XX forensic biology; cellular protease activity; cell interaction;

KW development; blood disorder; haematopoietic cell-related disorder;

KW growth; cell proliferation; cell differentiation; gamma delta T-cell;

KW immune system; metabolic disorder; homeostatic disorder; anaemia;

KW developmental bone disorder; osteoporosis; bacterial infection;

KW thrombocytopaenia; renal failure; liver disease; cystic fibrosis;

KW obesity; cancer; bronchitis; asthma; emphysema; pulmonary oedema;

KW respiratory distress syndrome; osteopathic; antibacterial; antianaemic;

KW thrombolytic; nephrotropic; antiobesity; hepatotropic; cytostatic;

KW antiinflammatory; antiasthmatic.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..33

FT /label= Predicted signal peptide

FT /note= "Given as SEQ ID No:48 and specifically

FT claimed in Claim 8"

FT Protein 34..423

FT /label= Predicted mature TANGO_294 protein

FT /note= "Given as SEQ ID No:49 and specifically

FT claimed in Claim 8"

FT Domain 34..254

FT /label= Extracellular domain

FT /note= "Given as SEQ ID No:50 and specifically

FT claimed in Claim 8"

FT Domain 255..279

FT /label= Transmembrane domain

FT /note= "Given as SEQ ID No:51 and specifically

FT claimed in Claim 8"

FT Domain 280..423

FT /label= Cytoplasmic domain

FT /note= "Given as SEQ ID No:52 and specifically

FT claimed in Claim 8"

XX US2002182675-A1.

XX 05-DEC-2002.

XX 25-OCT-2001; 2001US-0042431.

XX 14-JUN-1999; 98US-0333159.

XX 24-MAY-2000; 2000US-0578063.

XX (MILL-) MILLENNIUM PHARM INC.

XX McCarthy SA, Fraser CC, Sharp JD, Barnes TM;

XX WPI; 2003-328617/31.

XX N-PSDB; ABX94106.

XX New TANGO and INTERCEPT proteins, useful as modulating agents in

PT regulating a variety of cellular processes, in chromosome mapping, in

PT tissue typing, and in forensic biology -

XX Claim 8; Fig 6A-6C; 232pp; English.

XX The present invention relates to the isolation of novel TANGO or

CC INTERCEPT proteins, and the polynucleotide sequences encoding them.

CC The invention discloses sequences for human TANGO 202, TANGO 234,

CC TANGO 265, TANGO 273, TANGO 286, TANGO 294 and INTERCEPT 296. The

CC sequences for murine TANGO 202 and TANGO 273 are also provided. The

CC TANGO polypeptides and the polynucleotide sequences encoding them are

CC useful as modulating agents in regulating a variety of cellular

CC processes. The polynucleotide sequences are useful as primers or

CC hybridisation probes for the detection of nucleic acids encoding

CC TANGO polypeptides, chromosome mapping, tissue typing, and in

CC forensic biology. TANGO 202 can be used to diagnose, prevent or

CC treat disorders relating to aberrant cellular protease activity,

CC inappropriate interaction of cells with mediators, inappropriate

CC development, and blood and haematopoietic cell-related disorders.

CC TANGO 234 can be used to modulate growth, proliferation, survival,

CC differentiation, and activity of gamma delta T-cells. TANGO 265 can

CC be used to prevent, diagnose and treat disorders characterised by

aberrant organisation or development of a tissue or organ, and for modulating differentiation of cells of the immune system. TANGO 273 is useful for diagnosing, treating or preventing e.g. metabolic, homeostatic and developmental bone disorders (e.g. osteoporosis), and bacterial infection. TANGO 286 is useful for treating or preventing e.g. anaemia, thrombocytopenia, renal failure or liver disease. TANGO 294 is useful for treating or preventing e.g. cystic fibrosis or obesity. INTERCEPT 296 is useful for diagnosing, treating or preventing e.g. cancers, bronchitis, cystic fibrosis, asthma, emphysema, pulmonary oedema, or adult and infant respiratory distress syndromes. The present sequence represents human TANGO 294.

Query Match 100.0%; Score 2247; DB 24; Length 423;
Best Local Similarity 100.0%; Pred. No. 6.8e-220;
Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLETLSRWIVSHRMEMWLLILVAYMFQNNVSVHMTKAVDPFAFNISEIIHQGYPC 60
Db 1 MLETLSRWIVSHRMEMWLLILVAYMFQNNVSVHMTKAVDPFAFNISEIIHQGYPC 60

Qy 61 EYEVATDGYLSVNRIPRGLVOPKKTGSRPVLLQHGLVGGASWISNLPNNSLGFIL 120
Db 61 EYEVATDGYLSVNRIPRGLVOPKKTGSRPVLLQHGLVGGASWISNLPNNSLGFIL 120

Qy 121 ADAGFDVVMGNSRGNAWSRKHKLISIDQDEFWAFSDYEMARFDPVAVINFILQKTGQEKI 180
Db 121 ADAGFDVVMGNSRGNAWSRKHKLISIDQDEFWAFSDYEMARFDPVAVINFILQKTGQEKI 180

Qy 181 YVGVSGQTWGTAFSTMPDLAKIKMYFALAPIATVKHAKSPGTFKLLPDMKIGLFP 240
Db 181 YVGVSGQTWGTAFSTMPDLAKIKMYFALAPIATVKHAKSPGTFKLLPDMKIGLFP 240

Qy 241 GKXGFLYQTRFLRLQVILCOVILQDQCSNIMLLGFSNTNNMGRASYAHTLAGT 300
Db 241 GKXGFLYQTRFLRLQVILCOVILQDQCSNIMLLGFSNTNNMGRASYAHTLAGT 300

Qy 301 SVONILHWSQVNSGELRAFQWSETKLEKNCQPTPVYRVDMVTPTAMWTGGQDWLS 360
Db 301 SVONILHWSQVNSGELRAFQWSETKLEKNCQPTPVYRVDMVTPTAMWTGGQDWLS 360

Qy 361 NPEDVKMLSEVTNLIYHKNIPEWAVDFIWLGDAPHMYNEIHLMQQEBTNLSQGRCE 420
Db 361 NPEDVKMLSEVTNLIYHKNIPEWAVDFIWLGDAPHMYNEIHLMQQEBTNLSQGRCE 420

Qy 421 AVL 423
Db 421 AVL 423

RESULT 3
ID AAE11931 standard; Protein; 409 AA.
AC AAE11931;
XX AAE11931;
DT 18-DEC-2001 (first entry)
XX Human CG162 (or C59) lipase protein #2.
DE Human; apolipoprotein; lipase; lipid metabolism; myocardial infarction;
XX cardiovascular disease; arterial thrombosis; thrombolytic; antilipemic;
KW cerebral ischaemia; arterial thrombosis; thrombolytic; antilipemic;
KW coronary artery thrombosis; cerebral artery thrombosis; stroke;
KW intracardiac thrombosis; gene therapy; cardiovascular; vasodilator;
KW neuroprotectant; cerebroprotective.
XX Homo sapiens.
OS Homo sapiens.
XX Key Location/Qualifiers
FH Peptide 1..19
FT

Protein /label= Signal_peptide
20.409
/note= "Human mature CG162 (or C59) lipase protein"

WO200179446-A2.
25-OCT-2001.
16-APR-2001; 2001WO-US12529.
14-APR-2000; 2000US-197137P.
20-JUN-2000; 2000US-0598042.
03-AUG-2000; 2000US-0631451.
17-NOV-2000; 2000US-0667298.
17-NOV-2000; 2000US-0714936.
(HYSE-) HYSEQ INC.
PA Ballinger DG, Loeb D, Montgomery JR, Tang TY, Zhou P, Goodrich R;
PI Liu C, Asundi V, Zhao QA, Wehrman T, Drmanac RI, Ren F, Qian XB;
PI Wang D;
XX WPI; 2001-611724/70.
DR N-PSDB; AAD19226.
XX Nucleic acids encoding human apolipoproteins, lipases, and lipoprotein
PT receptor polypeptides, useful for preventing diagnosing and treating
PT lipid metabolism disorders, thrombosis and cardiovascular diseases -
XX Claim 10; Fig 3; 266pp; English.
XX The invention relates to polynucleotides encoding proteins CG122, CG179,
CC CG95, CG121, CG162, CG27, CG153 and CG168 which are related to proteins
CC involved in lipid metabolism and cardiovascular disease such as human
CC apolipoproteins, lipases and lipoprotein receptor proteins. These DNA
CC and protein sequences are useful for treating or preventing disorders
CC associated with apolipoproteins, lipases and lipoprotein receptor (ALLr)
CC expression and for treating lipid metabolism, cardiovascular diseases
CC and thrombosis. Antibodies against these proteins are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of these sequences. ALLr polypeptides are also
CC useful for identifying agents (agonists and antagonists) that bind to
CC them and cells expressing ALLr proteins are useful for identifying a
CC therapeutic agent for use in treatment of a pathology related to
CC aberrant expression or physiological interactions of this polypeptide.
CC Vectors comprising these DNA and protein sequences are also useful for
CC producing ALLr proteins. The nucleic acids and polypeptides of the
CC invention are also useful for the treatment of occlusive cardiovascular
CC diseases, myocardial infarction, cerebral ischaemia, angina, arterial
CC thrombosis, coronary artery thrombosis and cerebral artery thrombosis
CC or intracardiac thrombosis and stroke. The nucleotides of the invention
CC are used in gene therapy. The present sequence is human CG162 (or C59)
CC lipase protein.
XX
SQ Sequence 409 AA;
Query Match 96.8%; Score 2174; DB 22; Length 409;
Best Local Similarity 100.0%; Pred. No. 1.8e-212;
Matches 409; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 MEMWLLILVAYMFQNNVSVHMTKAVDPFAFNISEIIHQGYPCYEVATEGYILS 74
Db 1 MEMWLLILVAYMFQNNVSVHMTKAVDPFAFNISEIIHQGYPCYEVATEGYILS 60

Qy 75 VNRIPRGLVOPKKTGSRPVLLQHGLVGGASWISNLPNNSLGFILADAGFDVVMGNSRG 134
Db 61 VNRIPRGLVOPKKTGSRPVLLQHGLVGGASWISNLPNNSLGFILADAGFDVVMGNSRG 120

Qy 135 NAWSRKHKLISIDQDEFWAFSDYEMARFDPVAVINFILQKTGQEKIYYGVSGQTMGFI 194
Db 121 NAWSRKHKLISIDQDEFWAFSDYEMARFDPVAVINFILQKTGQEKIYYGVSGQTMGFI 180

Qy 195 AFSTMPDLAQKIMYFALAPIATVKHAKSPGTFKLLPDMKIGLFGKKEFLYQTRFLRQ 254

Db 181 AFSTMPBLAQKIMYFALAPIATVGHAKSPGTFKLLPDMTKGLFGKKEFLYQTRFLRQ 240
 Qy 255 LVLYLGGQVLDQICSNIMLLGGFTNNMNSRASVYAAHTLAGTSVQNIHLHWSQAVNS 314
 Db 241 LVLYLGGQVLDQICSNIMLLGGFTNNMNSRASVYAAHTLAGTSVQNIHLHWSQAVNS 300
 Qy 315 GELRAFDMGSETKXLEKCNQPTPVRYRVRDMTPTAMWTGGQDNLSPEDVVKLLSEVTN 374
 Db 301 GELRAFDMGSETKXLEKCNQPTPVRYRVRDMTPTAMWTGGQDNLSPEDVVKLLSEVTN 360
 Qy 375 LIYHKNIPWAHVDFIWLGDAPHRMYNEIHLMOQBEETNLSQRCBAVL 423
 Db 361 LIYHKNIPWAHVDFIWLGDAPHRMYNEIHLMOQBEETNLSQRCBAVL 409

RESULT 4
 AAB66067
 ID AAB66067 standard; Protein; 390 AA.
 XX
 AC AAB66067;
 XX
 DT 30-MAR-2001 (first entry)
 XX
 DE Human TANGO 294 mature protein.
 XX

XX TANGO protein; INTERCEPT protein; neurological disorder;
 KW central nervous system; focal brain disorder; bipolar affective disorder;
 KW global diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
 KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
 KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
 KW neuropsychiatric; psychoactive substance use; anxiety.
 XX

OS Homo sapiens.
 XX
 PN WO200077239-A2.
 XX
 PD 21-DEC-2000.
 XX

PF 24-MAY-2000; 2000WO-US14958.
 XX
 PR 14-JUN-1999; 99US-0333159.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX

PI McCarthy SA, Fraser CC, Sharp JD, Barnes TW;
 XX
 DR WPI; 2001-032313/04.
 XX

XX TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
 PT screening assays and diagnostic assays and for the treatment of
 PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
 PT disease -
 XX

PS Claim 8; Pages 324-325; 359pp; English.
 XX

CC The present invention relates to TANGO or INTERCEPT proteins and coding
 CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
 CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
 CC sequences are useful for the treatment of neurological disorders such as
 CC central nervous system (CNS) disorders, CNS-related disorders, focal
 CC brain disorders, global-diffuse cerebral disorders and other
 CC neurological and cerebrovascular disorders. The CNS disorders include
 CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
 CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
 CC autonomic function disorders such as hypertension and sleep disorders,
 CC neuropsychiatric disorders, psychoactive substance use disorders,
 CC anxiety, and bipolar affective disorder.
 XX

SQ Sequence 390 AA;
 Query Match 92.4%; Score 2076; DB 22; Length 390;
 Best Local Similarity 100.0%; Pred. No. 1.6e-202;

Matches 390; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 34 VHMPTKAVDPEAFNFIIEIIHQGYPCBEYEVATEDGYILSVNRI PRGLVQPKKTSRPV 93
 Db 1 VHMPTKAVDPEAFNFIIEIIHQGYPCBEYEVATEDGYILSVNRI PRGLVQPKKTSRPV 60
 Qy 94 VLLQHLVGGASNIWISNLPNNSLGFILADAGFDVWNGSRGNASRKHKTLSDQDEFWA 153
 Db 61 VLLQHLVGGASNIWISNLPNNSLGFILADAGFDVWNGSRGNASRKHKTLSDQDEFWA 120
 Qy 154 FSYDEMARFDPFPAVINFILOKTGOEKIYVYVGSQGTMTGFIAPSTMPBLAQKIMYFALA 213
 Db 121 FSYDEMARFDPFPAVINFILOKTGOEKIYVYVGSQGTMTGFIAPSTMPBLAQKIMYFALA 180
 Qy 214 PIATVGHAKSPGTFKLLPDMTKGLFGKKEFLYQTRFLRQVILYICGVLDQICSNIM 273
 Db 181 PIATVGHAKSPGTFKLLPDMTKGLFGKKEFLYQTRFLRQVILYICGVLDQICSNIM 240
 Qy 274 LLLGGFTNNMNSRASVYAAHTLAGTSVQNIHLHWSQAVNSGELRAFDMGSETKXLEKCN 333
 Db 241 LLLGGFTNNMNSRASVYAAHTLAGTSVQNIHLHWSQAVNSGELRAFDMGSETKXLEKCN 300
 Qy 334 QPTEPVRYRVRDMTPTAMWTGGQDNLSPEDVVKLLSEVTNLIYHKNIPWAHVDFIWL 393
 Db 301 QPTEPVRYRVRDMTPTAMWTGGQDNLSPEDVVKLLSEVTNLIYHKNIPWAHVDFIWL 360
 Qy 394 DAPHRMYNEIHLMOQBEETNLSQRCBAVL 423
 Db 361 DAPHRMYNEIHLMOQBEETNLSQRCBAVL 390

RESULT 5
 ID ABG31894 standard; Protein; 423 AA.
 XX
 AC ABG31894;
 XX
 DT 05-NOV-2002 (first entry)
 XX
 DE Human lipid-associated molecule, LIPAMI.
 XX
 KW Human; lipid-associated molecule; LIPAM; cancer; leukaemia;
 KW cardiovascular disorder; atherosclerosis; hypertension; stroke;
 KW angina pectoris; neurological disorder; Alzheimer's disease; epilepsy;
 KW dementia; Parkinson's disease; multiple sclerosis; anxiety; psoriasis;
 KW autoimmune disorder; inflammatory disorder; anaemia; asthma; trauma;
 KW acquired immunodeficiency syndrome; AIDS; bronchitis; diabetes mellitus;
 KW emphysema; Crohn's disease; atopic dermatitis; rheumatoid arthritis;
 KW gastrointestinal disorder; lipid metabolism disorder; transgenic animal.

OS Homo sapiens.
 XX
 PN WO200246418-A2.
 XX
 PD 13-JUN-2002.
 XX

PF 04-DEC-2001; 2001WO-US47430.
 XX
 PR 08-DEC-2000; 2000US-254505P.
 PR 15-DEC-2000; 2000US-256187P.
 PR 22-DEC-2000; 2000US-257908P.
 PR 26-JAN-2001; 2001US-264429P.
 XX (INCY-) INCYTE GENOMICS INC.
 PA Griffin JA, Gandhi AR, Ramkumar J, Tang YT, Ding L, Yue H;
 PI Gietzen KJ, Sapperstein SK, Honchell CD, Bruns CM, Duggan BM;
 PI Xu Y, Lee S;
 XX

DR WPI; 2002-619022/66.
 DR N-PSDB; ABR90871.
 XX
 PT Novel polypeptide of human lipid associated molecule, useful for

PT diagnosing, treating and preventing cancer, hypercholesterolaemia,
PT cirrhosis, myocardial infarction, Parkinson's disease, asthma,
XX psoriasis, gastritis
PS Claim 1; Page 111-112; 126pp; English.
XX
CC The invention relates to an isolated polypeptide (I) of human lipid-
CC associated molecule (LIPAM), and an isolated polynucleotide (II) encoding
CC (I). (I) and antibody to (II) are useful for treating a disease or
CC condition associated with decreased or increased expression of functional
CC LIPAM. (I) or (II) is useful for diagnosing, treating or preventing
CC cancer e.g. leukemia, cancer of adrenal gland, bladder, bone, bone
CC marrow, brain, ovary, etc; cardiovascular disorders e.g. atherosclerosis,
CC hypertension, Raynaud's disease, angina pectoris, myocardial infarction,
CC rheumatic fever, cardiomyopathy, pericarditis, chronic obstructive
CC pulmonary disease, etc; neurological disorders e.g. Alzheimer's disease,
CC stroke, epilepsy, dementia, Parkinson's disease, multiple sclerosis,
CC Creutzfeldt-Jakob disease, myasthenia gravis, anxiety, diabetic
CC neuropathy, etc; autoimmune/inflammatory disorders e.g. acquired
CC immunodeficiency syndrome (AIDS), Addison's disease, anaemia, asthma,
CC bronchitis, diabetes mellitus, emphysema, Crohn's disease, atopic
CC dermatitis, rheumatoid arthritis, psoriasis, systemic lupus
CC erythematosus, trauma, etc; and gastrointestinal disorders e.g.
CC gastritis, anorexia, nausea gastroenteritis, ulcerative colitis,
CC cholecystitis, hepatitis, cirrhosis, hepatoma, diarrhoea, constipation,
CC etc, and disorders of lipid metabolism e.g., fatty liver, cholestasis,
CC Fabry's disease, hypercholesterolaemia. (I) or (II) is useful for
CC assessing the effects of exogenous compounds on the expression of (I)
CC or (II). (I) is useful in number of drug screening techniques, and to
CC analyse the proteome of a tissue or cell type. (I) or antibody to (II) is
CC useful as elements on a microarray. (II) is useful for creating knockin
CC humanized animals or transgenic animals to model human diseases, in
CC somatic or germline gene therapy, to generate a transcript image, of a
CC tissue or cell type, for detecting differences in the chromosomal
CC location due to translocation, inversion, etc., among normal, carrier or
CC affected individuals, and as hybridisation probes for mapping naturally
CC occurring genomic sequences. The present sequence represents the amino
CC acid sequence of a human lipid-associated molecule (LIPAM) of the
CC invention.
XX
SQ Sequence 423 AA;
Query Match 89.6%; Score 2013.5; DB 23; Length 423;
Best Local Similarity 92.1%; Pred. No. 4.2e-196;
Matches 384; Conservative 9; Mismatches 13; Indels 11; Gaps 3;
QY 17 MWLLILVAYMFORVNS---VHPTKAVDPDAFNISIIHQHQPCEYEVEATEDGYIL 73
DB 8 MWLLITTTCLIGTLAGGFLDLENE-VNPEVMMNTSIIHQHQPCEYEVEATEDGYIL 66
QY 74 SVNRIPLGLVOPKKTGSRPVLLQHLVGGASNIISLNNPNSLGFILADAGFDVWGNRS 133
DB 67 SVNRIPLGLVOPKKTGSRPVLLQHLVGGASNIISLNNPNSLGFILADAGFDVWGNRS 126
QY 134 GNWSRKHTLSIDQDEFWAFSYDEMARFDPVAVINFILQKTQGEKIVYVYSGQTTMGF 193
DB 127 GNWSRKHTLSIDQDEFWAFSYDEMARFDPVAVINFILQKTQGEKIVYVYSGQTTMGF 186
QY 194 IAFSTMPPELAOKIKMYPALAPATVTKHAKSGPTKFLPLPDMWIKGLGKKEFLYQTRFLR 253
DB 187 IAFSTMPPELAOKIKMYPALAPATVTKHAKSGPTKFLPLPDMWIKGLGKKEFLYQTRFLR 246
QY 254 QLVYLCQGVILQDQCSNIMLLGFGNTNNNM-----SRASVYAHTLAGTSVQNIL 306
DB 247 QLVYLCQGVILQDQCSNIMLLGFGNTNNNMTHGLLQSRASVYAHTLAGTSVQNIL 306
QY 307 HWSQAVNSGEARAFDWGSETNLEKCNQPTVPRVRVDMTPTAMWTGGQDLNSPEDVK 366
DB 307 HWSQAVNSGEARAFDWGSETNLEKCNQPTVPRVRVDMTPTAMWTGGQDLNSPEDVK 366
QY 367 MLLSEVTNLIYHKNIPEWAHVDIFWGLDAPHRMYNEIHLMQQBEETNLSQRCCEAVL 423
DB 367 MLLSEVTNLIYHKNIPEWAHVDIFWGLDAPHRMYNEIHLMQQBEETNLSQRCCEAVL 423

RESULT 6
ABG71000
ID ABG71000 standard; Protein; 398 AA.
XX
AC ABG71000;
XX
XX 07-JAN-2003 (first entry)
XX
DE Human lipase protein.
XX
XX Human; lipase; cancer; colon; kidney; skin; brain; testis; chromosome 10;
XX Burkitt's lymphoma; cytostatic; gene therapy; lipase-inhibitor.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FT Modified-site 88..91 /note= "N-glycosylation site"
FT Modified-site 260..263 /note= "N-glycosylation site"
FT Modified-site 388..391 /note= "N-glycosylation site"
FT Modified-site 22..25 /note= "cAMP- and cGMP-dependent protein kinase phosphorylation site"
FT Modified-site 113..115 /note= "Protein kinase C phosphorylation site"
FT Modified-site 192..194 /note= "Protein kinase C phosphorylation site"
FT Modified-site 130..133 /note= "Casein kinase II phosphorylation site"
FT Modified-site 150..153 /note= "Casein kinase II phosphorylation site"
FT Modified-site 173..176 /note= "Casein kinase II phosphorylation site"
FT Modified-site 335..338 /note= "Casein kinase II phosphorylation site"
FT Modified-site 149..157 /note= "Casein kinase II phosphorylation site"
FT Modified-site 74..79 /note= "Tyrosine kinase phosphorylation site"
FT Modified-site 77..82 /note= "N-myristoylation site"
FT Modified-site 105..110 /note= "N-myristoylation site"
FT Modified-site 109..114 /note= "N-myristoylation site"
FT Modified-site 163..168 /note= "N-myristoylation site"
FT Modified-site 252..257 /note= "N-myristoylation site"
FT Modified-site 215..218 /note= "Amidation site"
FT Active-site 155..164 /note= "Lipases, serine active site"
XX
XX WO200274975-A2.
XX
XX 26-SEP-2002.
XX
XX 18-MAR-2002; 2002WO-US08035.
XX
XX 20-MAR-2001; 2001US-0811825.
XX
XX (PEKE) PE CORP NY.
XX
XX Yan C, Di Francesco V, Beasley EM;
XX WPI; 2002-750560/81.
XX N-PSDB; ABS55356, ABS55357.
XX
XX

PT New isolated human lipase peptides and encoding nucleic acids, useful
PT for diagnosing and treating disorders mediated by human lipase
PT proteins, e.g. cancer of the colon, kidneys, skin and testis, and
PT Burkitt's lymphoma -

XX Claim 1; Fig 2; 79pp; English.

CC The present invention relates to a new lipase peptide. The methods and
CC compositions of the present invention are useful for diagnosing and
CC treating disorders mediated by the human lipase protein, such as cancer
CC of the colon, kidneys, skin, brain and testis, and Burkitt's lymphoma.
CC The present amino acid sequence represents the human lipase protein of
CC the invention. This sequence is encoded by the human lipase gene
CC located on chromosome 10.

XX SQ Sequence 398 AA;

Query Match 88.6%; Score 1991; DB 23; Length 398;

Best Local Similarity 100.0%; Pred. No. 7.6e-194;

Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 SEIIHQGYPCPEYEVATEDGYILSVNRI PRGLVQPKTGSRPVLLQHLGVGASNWIS 109

Db 25 SEIIHQGYPCPEYEVATEDGYILSVNRI PRGLVQPKTGSRPVLLQHLGVGASNWIS 84

QY 110 NLPNNSLGFILADAGFDVWNGSRGNWSKHKHTLSIDQDEFWASDEMARFDPVAVIN 169

Db 85 NLPNNSLGFILADAGFDVWNGSRGNWSKHKHTLSIDQDEFWASDEMARFDPVAVIN 144

QY 170 FILQKTQGEKIYYVYSGQTTMGFIAPSTMPPELAQIKMYFALAPIATVHKASPGTKFL 229

Db 145 FILQKTQGEKIYYVYSGQTTMGFIAPSTMPPELAQIKMYFALAPIATVHKASPGTKFL 204

QY 230 LLPDMNKGFLGKKEFLYQTRFLRQVLYLCGVILDOICSNIMLLGGFTNNMNSRA 289

Db 205 LLPDMNKGFLGKKEFLYQTRFLRQVLYLCGVILDOICSNIMLLGGFTNNMNSRA 264

QY 290 SVYAAHTLAGTSVQNILHWSQAVNSGELRAFDMGSETKLEKNCQPTPVRYVRDMTVPT 349

Db 265 SVYAAHTLAGTSVQNILHWSQAVNSGELRAFDMGSETKLEKNCQPTPVRYVRDMTVPT 324

QY 350 AMWTGGQDWLSNPEDVKMLSEVNTLIYHKNIPEWAHVDFIWLGDAPHMYNEIHLMOQ 409

Db 325 AMWTGGQDWLSNPEDVKMLSEVNTLIYHKNIPEWAHVDFIWLGDAPHMYNEIHLMOQ 384

QY 410 EETNLSQGRCEAVL 423

Db 385 EETNLSQGRCEAVL 398

RESULT 7

AAU99164

XX ID AAU99164 standard; Protein; 398 AA.

XX NC AAU99164;

XX DT 24-SEP-2002 (first entry)

XX DE Human lysosomal acid lipase #1.

XX KW Human; enzyme; lysosomal acid lipase; lipid malabsorption illness;

XX KW cystic fibrosis; alcoholism; heart disease; heart attack;

XX KW Wolman disease; cholesterol ester storage disease; brain injury;

XX KW mood disorder; anxiety disorder; thought disorder; volition disorder;

XX KW sleep disorder; neurogenic disorder; myopathic disorder; COPD;

XX KW obesity; cancer; chronic obstructive pulmonary disease; diabetes;

XX KW cardiovascular disorder; Alzheimer's disease; Parkinson's disease;

XX KW anorexia; osteoarthritis; central nervous system disorder;

XX KW peripheral nervous system disorder.

XX OS Homo sapiens.

XX XX

XX FN WO200236731-A2.

XX 10-MAY-2002.

XX PD 30-OCT-2001; 2001WO-EPI2518.

XX PF 31-OCT-2000; 2000US-244170P.

XX PR 29-MAY-2001; 2001US-293516P.

XX PA (FARB) BAYER AG.

XX PI Xiao Y;

XX XX WPI; 2002-519248/55.

XX DR N-PSDB; ABK86569.

XX XX Novel human lysosomal acid lipase polypeptide, useful for treating

PT cancer, diabetes, obesity, chronic obstructive pulmonary disease,

PT peripheral or central nervous system disorder or cardiovascular

PT disorder -

XX Claim 25; Fig 2; 126pp; English.

XX The invention relates to a purified human lysosomal acid lipase

CC polypeptide. Also included are the polynucleotide encoding the

CC lipase (or its fragment, derivative, allele or sequence at least 60%

CC identical to it), vectors, host cells, a reagent (e.g. an antisense

CC oligonucleotide), which binds to the lipase or polynucleotide (used for

CC detection and modulating/reducing the lipase activity) and an anti-lipase

CC antibody. The lipase and polynucleotide are useful for identifying

CC therapeutic agents that either increase or decrease the lipase activity.

CC The identified agent, the lipase and polynucleotide are useful for

CC treatment of a disease such as lipid malabsorption illness,

CC cystic fibrosis, alcoholism, heart disease, heart attack,

CC Wolman disease, cholesterol ester storage disease, brain injury,

CC mood disorder, anxiety disorder, thought disorder, volition disorder,

CC sleep disorder, neurogenic disorder, myopathic disorder, obesity, cancer,

CC chronic obstructive pulmonary disease (COPD), diabetes,

CC cardiovascular disorder, Alzheimer's disease, Parkinson's disease,

CC anorexia, osteoarthritis, a central nervous system disorder and

CC a peripheral nervous system disorder. The present sequence is the

CC human lysosomal lipase #1.

XX SQ Sequence 398 AA;

Query Match 88.6%; Score 1991; DB 23; Length 398;

Best Local Similarity 100.0%; Pred. No. 7.6e-194;

Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 SEIIHQGYPCPEYEVATEDGYILSVNRI PRGLVQPKTGSRPVLLQHLGVGASNWIS 109

Db 25 SEIIHQGYPCPEYEVATEDGYILSVNRI PRGLVQPKTGSRPVLLQHLGVGASNWIS 84

QY 110 NLPNNSLGFILADAGFDVWNGSRGNWSKHKHTLSIDQDEFWASDEMARFDPVAVIN 169

Db 85 NLPNNSLGFILADAGFDVWNGSRGNWSKHKHTLSIDQDEFWASDEMARFDPVAVIN 144

QY 170 FILQKTQGEKIYYVYSGQTTMGFIAPSTMPPELAQIKMYFALAPIATVHKASPGTKFL 229

Db 145 FILQKTQGEKIYYVYSGQTTMGFIAPSTMPPELAQIKMYFALAPIATVHKASPGTKFL 204

QY 230 LLPDMNKGFLGKKEFLYQTRFLRQVLYLCGVILDOICSNIMLLGGFTNNMNSRA 289

Db 205 LLPDMNKGFLGKKEFLYQTRFLRQVLYLCGVILDOICSNIMLLGGFTNNMNSRA 264

QY 290 SVYAAHTLAGTSVQNILHWSQAVNSGELRAFDMGSETKLEKNCQPTPVRYVRDMTVPT 349

Db 265 SVYAAHTLAGTSVQNILHWSQAVNSGELRAFDMGSETKLEKNCQPTPVRYVRDMTVPT 324

QY 350 AMWTGGQDWLSNPEDVKMLSEVNTLIYHKNIPEWAHVDFIWLGDAPHMYNEIHLMOQ 409

Db 325 AMWTGGQDWLSNPEDVKMLSEVNTLIYHKNIPEWAHVDFIWLGDAPHMYNEIHLMOQ 384

QY 410 EETNLSQGRCEAVL 423

Db 385 EETNLSQRCFAVL 398

RESULT 8

AAU99165 ID AAU99165 standard; Protein; 371 AA.

XX AC AAU99165;

XX DT 24-SEP-2002 (first entry)

XX DE Human lysosomal acid lipase #2.

XX KW Human; enzyme; lysosomal acid lipase; lipid malabsorption illness; cystic fibrosis; alcoholism; heart disease; heart attack; Wolman disease; cholesterol ester storage disease; brain injury; mood disorder; anxiety disorder; thought disorder; volition disorder; sleep disorder; neurogenic disorder; myopathic disorder; COPD; obesity; cancer; chronic obstructive pulmonary disease; diabetes; cardiovascular disorder; Alzheimer's disease; Parkinson's disease; anorexia; osteoarthritis; central nervous system disorder; peripheral nervous system disorder.

XX OS Homo sapiens.

XX FN WO200236731-A2.

XX PD 10-MAY-2002.

XX PP 30-OCT-2001; 2001WO-EP12518.

XX PR 31-OCT-2000; 2000US-244170P.

XX PR 29-MAY-2001; 2001US-293516P.

XX PA (FARB) BAYER AG.

XX PI Xiao Y;

XX DR WPI; 2002-519248/55.

XX DR N-PSDB; ABK86570.

XX PT Novel human lysosomal acid lipase polypeptide, useful for treating cancer, diabetes, obesity, chronic obstructive pulmonary disease, peripheral or central nervous system disorder or cardiovascular disorder.

XX PS Claim 25; Fig 5; 126pp; English.

XX CC The invention relates to a purified human lysosomal acid lipase polypeptide. Also included are the polynucleotide encoding the lipase (or its fragment, derivative, allele or sequence at least 60% identical to it), vectors, host cells, a reagent (e.g. an antisense oligonucleotide) which binds to the lipase or polynucleotide (used for detection and modulating/reducing the lipase activity) and an anti-lipase antibody. The lipase and polynucleotide are useful for identifying therapeutic agents that either increase or decrease the lipase activity. The identified agent, the lipase and polynucleotide are useful for treatment of a disease such as lipid malabsorption illness, cystic fibrosis, alcoholism, heart disease, heart attack, Wolman disease, cholesterol ester storage disease, brain injury, mood disorder, anxiety disorder, thought disorder, volition disorder, sleep disorder, neurogenic disorder, myopathic disorder, obesity, cancer, chronic obstructive pulmonary disease (COPD), diabetes, cardiovascular disorder, Alzheimer's disease, Parkinson's disease, anorexia, osteoarthritis, a central nervous system disorder and a peripheral nervous system disorder. The present sequence is the human lysosomal lipase #2.

XX SQ Sequence 371 AA;

Query Match 68.4%; Score 1536.5; DB 23; Length 371;
Best Local Similarity 80.9%; Pred. No. 1.4e-147;

	Matches	297;	Conservative	17;	Mismatches	22;	Indels	31;	Gaps	2;
QY	50	SEIIHQGYPC	EEYEVATEDGYILSVNRI	PRGLVQPKTKGSRPVVLLQHGVLGGASNWIS	109					
Db	25	SEIIHQGYPC	EEYEVATEDGYILSVNRI	PRGLVQPKTKGSRPVVLLQHGVLGGASNWIS	84					
QY	110	NLPNNSLGL	FIILADAGFDVVMGNSRGNWSRKHKTLSDQDEFWAFSDYDEMARFDP	PAVIN	169					
Db	85	NLPNNSLGL	FIILADAGFDVVMGNSRGNWSRKHKTLSDQDEFWAFSDYDEMARFDP	PAVIN	144					
QY	170	FILQKTGQEK	IIYVGYSGTGTMGFTAFSTMPBELAQIKMYFALAPIATVVKHAKSPG	KFL	229					
Db	145	FILQKTGQEK	IIYVGYSGTGTMGFTAFSTMPBELAQIKMYFALAPIATVVKHAKSPG	KFL	204					
QY	230	LIPDMWIKGL	FGKKEFLYQTRFLRQVLYLQGVILDOICSNIMLLGGFNTNNNM	---	286					
Db	205	LIPDMWIKGL	FGKKEFLYQTRFLRQVLYLQGVILDOICSNIMLLGGFNTNNNM	WTH	264					
QY	287	----	SRASYAAHTLAGTSVONILHWSQAVNSGELRAFDWGSSEKKNLEKQNP	TPVRYV	342					
Db	265	GLQSRASYA	AAHTLAGTSVONILHWSQ-----	-----TPPLYNI	300					
QY	343	RDMTPTAMWT	GGQDWLSNPEDVKMLLSEVTNLIYHKNIPEWAHVDFRWGLDAPHERMYNE	402						
Db	301	TKIEVPTAI	WGGQDIVADPKDVENLLPQIANLIYYKLIPHYNHVDVFLGEDAPQEII	YQD	360					
QY	403	IIHLMQ	409							
Db	361	LIILMEE	367							

RESULT 9

AAU90783 ID AAB90783 standard; Protein; 399 AA.

XX AC AAB90783;

XX DT 15-JUN-2001 (first entry)

XX DE Human shear stress-response protein SEQ ID NO: 66.

XX KW Human; shear stress-response protein; vascular disease; arteriosclerosis.

XX OS Homo sapiens.

XX FN WO200125427-A1.

XX PD 12-APR-2001.

XX PF 02-OCT-2000; 2000WO-JP06840.

XX PR 01-OCT-1999; 99JP-0280976.

XX PA (KYOW) KYOWA HAKKO KOGYO KK.

XX PI (NOJI/) NOJIMA H.

XX PI Nojima H, Yoshisue H, Obayashi M, Ota T, Kawabata A, Sakurada K, Kuga T, Sekine S, Nakamura Y, Sugano S;
XX WPI; 2001-266308/27.
XX DR N-PSDB; AAH02906.
XX PT DNA sequences, proteins encoded by them and antibodies against them useful in diagnosis and treatment of vascular disease caused by arteriosclerosis -
XX PS Claim 60; Page 402-404; 678pp; Japanese.

XX CC The present invention provides the protein and coding sequences of a number of human shear stress response proteins. These are useful in the diagnosis, treatment and screening of vascular diseases caused by arteriosclerosis, including heart failure, post-PTCA restenosis and

CC hypertension.
XX
SQ Sequence 399 AA;

Query Match 57.4%; Score 1289; DB 22; Length 399;
Best Local Similarity 63.5%; Pred. No. 2.6e-122;
Matches 235; Conservative 55; Mismatches 80; Indels 0; Gaps 0;

QY 40 AVDPEAFNMISIIHQGYPCPEYEVATEDGYILSVNRIIPRGLVQPKTGRPVVLLQHG 99
DB 28 AVDPETNMVSEIIISYWGPPSEYLVETEDGYILCLNRIIPRGRKNHSDKGPVVFLLQHG 87

QY 100 LVGASNWTISNPNNSLGFILADAGFDVVMGNSRGNSNRKHKTLISIDQDEFWAFSYDEM 159
DB 88 LLADSNWNTNLANSLSLGFILADAGFDVVMGNSRGNTWSRKHKTLISVSDQDEFWAFSYDEM 147

QY 160 ARFDLPAVINFLQKTQGBKIIYVYGSQGTMGFIAPSTMPBLAQKIMYFALAPIATVK 219
DB 148 AKYDLPASINFILNKTQGBQVYVGHSGQTTIGFIAPSOIPELAKRIKMFALGPVASVA 207

QY 220 HAKSPGTEKLLPDMMIKGLFGKKEFLYQTRFLQVILYLCQVILDOICSNIMLLGGF 279
DB 208 FCTSPMAKLGRLPDHLIKDLFGDKFELPOSALFKMLGTHVTHVILKELCGNLCFLCGF 267

QY 280 NTNNMNSRASVYAAHTLAGTSVQNILHWSQAVNSGELRAFDWGSSETKNLEKCNQPTPVR 339
DB 268 NEENLANSRVDDVYTHSPAGTSVQNMHLWSQAVKQKQAFDWSGSAKNYFHYNQSYPT 327

QY 340 YRVDMTPTAMWTGGQDWSNPEDVXKMLSEVTLIYHKNIPEWAHVDFIWGLDAPHRM 399
DB 328 YNVKMLVPTAVWSGGHDLADVDVYNILLTQITNLVFHESIPWEHLDFIWGLDAPHRM 387

QY 400 YNEIHLMOQ 409
DB 388 YNKIINLMEK 397

RESULT 11
AAO18227
ID AAO18227 standard; Protein; 392 AA.
AC AAO18227;
XX
XX
DT 18-SEP-2002 (first entry)
DE Human lysosomal acid lipase related protein.
XX Human; lysosomal acid lipase; chromosome 10; lipoemic; Wolman disease;
KW cholesteryl ester storage disease.
XX Homo sapiens.
XX US6387680-B1.
XX PD 14-MAY-2002.
XX

CC hypertension.
XX
SQ Sequence 399 AA;

Query Match 57.4%; Score 1289; DB 22; Length 399;
Best Local Similarity 63.5%; Pred. No. 2.6e-122;
Matches 235; Conservative 55; Mismatches 80; Indels 0; Gaps 0;

QY 40 AVDPEAFNMISIIHQGYPCPEYEVATEDGYILSVNRIIPRGLVQPKTGRPVVLLQHG 99
DB 28 AVDPETNMVSEIIISYWGPPSEYLVETEDGYILCLNRIIPRGRKNHSDKGPVVFLLQHG 87

QY 100 LVGASNWTISNPNNSLGFILADAGFDVVMGNSRGNSNRKHKTLISIDQDEFWAFSYDEM 159
DB 88 LLADSNWNTNLANSLSLGFILADAGFDVVMGNSRGNTWSRKHKTLISVSDQDEFWAFSYDEM 147

QY 160 ARFDLPAVINFLQKTQGBKIIYVYGSQGTMGFIAPSTMPBLAQKIMYFALAPIATVK 219
DB 148 AKYDLPASINFILNKTQGBQVYVGHSGQTTIGFIAPSOIPELAKRIKMFALGPVASVA 207

QY 220 HAKSPGTEKLLPDMMIKGLFGKKEFLYQTRFLQVILYLCQVILDOICSNIMLLGGF 279
DB 208 FCTSPMAKLGRLPDHLIKDLFGDKFELPOSALFKMLGTHVTHVILKELCGNLCFLCGF 267

QY 280 NTNNMNSRASVYAAHTLAGTSVQNILHWSQAVNSGELRAFDWGSSETKNLEKCNQPTPVR 339
DB 268 NEENLANSRVDDVYTHSPAGTSVQNMHLWSQAVKQKQAFDWSGSAKNYFHYNQSYPT 327

QY 340 YRVDMTPTAMWTGGQDWSNPEDVXKMLSEVTLIYHKNIPEWAHVDFIWGLDAPHRM 399
DB 328 YNVKMLVPTAVWSGGHDLADVDVYNILLTQITNLVFHESIPWEHLDFIWGLDAPHRM 387

QY 400 YNEIHLMOQ 409
DB 388 YNKIINLMEK 397

RESULT 10
AAB66061
ID AAB66061 standard; Protein; 399 AA.
AC AAB66061;
XX
XX
DT 30-MAR-2001 (first entry)
DE Human lysosomal acid lipase protein.
XX
XX
KW Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
KW central nervous system; focal brain disorder; bipolar affective disorder;
KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
KW neuropsychiatric; psychoactive substance use; anxiety.
XX
XX Homo sapiens.
XX
XX WO200077239-A2.
XX
XX 21-DEC-2000.
XX
XX 24-MAY-2000; 2000WO-US14858.
XX
XX 14-JUN-1999; 99US-0333159.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX McCarthy SA, Fraser CC, Sharp JD, Barnes TM;
XX
XX WPI; 2001-032313/04.
XX
XX TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
PT screening assays and diagnostic assays and for the treatment of
PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's

Db 407 LQITWLVFHEISPEWEHLDFIWLGLDAPWRLYNKLIINLWRK 447

RESULT 13

AAR37302
ID AAR37302 standard; Protein; 398 AA.

XX AAR37302;

XX 25-MAR-2003 (updated)

DT 20-SEP-1993 (first entry)

XX RGL precursor.

XX Rabbit gastric lipase; RGL; prGLN2.1; fat; bioconversion;

KW hydrolysis; transesterification.

XX Oryctolagus cuniculus.

OS

XX Key Location/Qualifiers
FT Protein 23..398
FT /note= "claim 1; page 10-11"

PN EP542629-A1.

XX 19-MAY-1993.

XX 12-NOV-1992; 92EP-0403055.

XX 13-NOV-1991; 91FR-0013948.

XX (LJOU) INST RECH JOUVEINAL.

XX Benicourt C, Blanchard C, Junien J;

XX WPI; 1993-161080/20.

XX N-PSDB; AAQ42310.

XX Rabbit gastric lipase, its precursor and their DNA - useful for

PT treating conditions linked to gastric lipase deficiency, such as

PT mucoviscidiosis and pancreatic exocrine insufficiency

XX Claim 1; Fig 7; 31pp; French.

XX RGL, opt. used with other lipases, are useful therapeutically

CC (1) to facilitate absorption of ingested fats in patients deficient

CC in endogenous gastric lipase and (2) to treat disorders caused by

CC inadequate lipase prodn. esp. mucoviscidiosis or pancreatic exocrine

CC deficiency. It can also be used (partic. when immobilised) for

CC enzymatic bioconversion, e.g. hydrolysis or transesterification.

CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 398 AA;

XX Query Match 51.9%; Score 1166; DB 14; Length 398;

XX Best Local Similarity 54.6%; Pred. No. 8.8e-110;

XX Matches 219; Conservative 70; Mismatches 102; Indels 10; Gaps 3;

QY 17 MWLLLVAYMFQNVNSVH-----WPTKAVDPEAFNNTSEIIIOHOGYCEEEVATEDG 70

Db 1 MWVLFMVAALISA-LGTHGLFGKSAPT---NPEVNNISQMSYWGYPSEKYEVVTEG 56

QY 71 YILSVNRIPRGLVQPKTKGSRPVVLLQHGLVGGASNWSLNPNSLGFILADAGFDVWVG 130

Db 57 YILEVNRIPYCKKSGNRGQRPVFLQHGLLASASNWSLNPNSLAFILADAGYGVWLG 116

QY 131 NSRGNSNRKHKLISIQDDEFAFSDYEMARFLPAVINFILOKTQEKLYYVGSQGT 190

Db 117 NSRGNTWSRNLYYSPDSVFPAFSDYEMAKYDLPATIDFIVKETGCKLHYVGSQGT 176

QY 191 MGFAFSTMPELAQIKWYFALAPIATVYHAKSPGTKFLLLPDMWIKGLFGKKEFLYQTR 250

Db 177 IGFIAFSTNPLAKRIKTFALAPVAIVTKYKSLVNNKLRFPPTMFKIFGDKIFYPHNF 236

QY 251 FLRLVYLGQVILDOICSNIMLLGGFNTNNMNSRASVYAAHTLACTSVQNIILHWSQ 310

Db 237 FDFLATQVCSRETINVICSNALFIICGFDNSANLMSRLDVTYVSHNPACTSVQNMHLWTQ 296

QY 311 AVNSGELRAPDWGSETKNLEKNQPPVRYRVRDMVTPTAMMTGGODWJUSNEDVKOLLUS 370

Db 297 AVLSGNFQAFNMGSPAQNVVHFNQPTPPYYNTYTMNVPIAVWSGGNDWLADPQVDLLLP 356

QY 371 EVTNLYHKNIPEWAHVDFRTWGLDAPHRMYNLIHLMQOEE 411

Db 357 KLSNLYHKELIPYNHLDPIWNNAPQEVNLIISMAKDK 397

RESULT 14

AAB66068

ID AAB66068 standard; Protein; 221 AA.

XX AAB66068;

XX 30-MAR-2001 (first entry)

XX Human TANGO 294 extracellular domain.

XX TANGO protein; INTERCEPT protein; neurological disorder;

KW central nervous system; focal brain disorder; bipolar affective disorder;

KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;

KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;

KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;

KW neuropsychiatric; psychoactive substance use; anxiety.

XX Homo sapiens.

XX WO2000077239-A2.

XX 21-DEC-2000.

XX 24-MAY-2000; 2000WO-US14858.

XX 14-JUN-1999; 99US-0333159.

XX (MILL-) MILLENNIUM PHARM INC.

XX McCarthy SA, Fraser CC, Sharp JD, Barnes TM;

XX WPI; 2001-032313/04.

XX TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for

PT screening assays and diagnostic assays and for the treatment of

PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's

PT disease -

XX Claim 8; Page 326; 359pp; English.

XX The present invention relates to TANGO or INTERCEPT proteins and coding

CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,

CC AAB66054-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding

CC sequences are useful for the treatment of neurological disorders such as

CC central nervous system (CNS) disorders, CNS-related disorders, focal

CC brain disorders, global-diffuse cerebral disorders and other

CC neurological and cerebrovascular disorders. The CNS disorders include

CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic

CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,

CC autonomic function disorders such as hypertension and sleep disorders,

CC neuropsychiatric disorders, psychoactive substance use disorders,

CC anxiety, and bipolar affective disorder.

XX Sequence 221 AA;

XX Query Match 51.8%; Score 1165; DB 22; Length 221;

XX Best Local Similarity 100.0%; Pred. No. 4.5e-110;

XX Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 VHMPTKAVDPEAFMNISEIIHQGYPCPEYEVATEDGYILSVNRIPRGLVQPKTGSRPV 93
 Db 1 VHMPTKAVDPEAFMNISEIIHQGYPCPEYEVATEDGYILSVNRIPRGLVQPKTGSRPV 60
 QY 94 VILQGLVGGASWISNLPNNISLGFILADAGFDVWGNISRGNAWSRKHTLSIDODEFWA 153
 Db 61 VILQGLVGGASWISNLPNNISLGFILADAGFDVWGNISRGNAWSRKHTLSIDODEFWA 120
 QY 154 FSYDEMARFDLPAVINFILOKQGOEKIYVGVYSGQTTWGFIAFSTMPPELAQIKMYFALA 213
 Db 121 FSYDEMARFDLPAVINFILOKQGOEKIYVGVYSGQTTWGFIAFSTMPPELAQIKMYFALA 180
 QY 214 PIATVHKASPGTKPFLLLPDMNKGFLGKKEFLYOTRFLRQ 254
 Db 181 PIATVHKASPGTKPFLLLPDMNKGFLGKKEFLYOTRFLRQ 221

RESULT 15

AA050322
 ID AAP50322 standard; protein; 395 AA.

XX AC AAP50322;
 XX AC
 DT 25-MAR-2003 (updated)
 DT 17-JAN-1992 (first entry)
 XX DT
 DE Rat lingual lipase protein.
 XX DE
 KW Lingual lipase; enzyme; EC-3.1.1.3; ss.
 XX KW
 OS Rattus rattus.
 XX OS
 PN GB2142337-A.
 XX PN
 PD 16-JAN-1985.
 XX PD
 PF 29-JUN-1984; 84GB-0016581.
 XX PF
 PR 29-JUN-1984; 84GB-0016581.
 XX PR
 XX (CLLT) CELLTech LTD.
 PA (WILL/) WILLIAMSON R.
 XX PA
 PI Carey NH, Williamson R;
 XX PI
 DR WPI; 1985-014450/03.
 DR N-PSDB; AAN50385.
 XX DR
 PT New lingual lipase protein for treatment of lipase deficiency - also
 PT new pre-lingual lipase protein and related products
 XX PT
 PS Disclosure; Fig 5; 15pp; English.
 XX PS
 CC This protein may be expressed in a transformant host organism
 CC and may be used for the treatment of lipase deficiency.
 CC (Updated on 25-MAR-2003 to correct PR field.)
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX CC
 SQ Sequence 395 AA;

Query Match 51.5%; Score 1156.5; DB 6; Length 395;
 Best Local Similarity 54.7%; Pred. No. 8.1e-109;
 Matches 216; Conservative 69; Mismatches 109; Indels 1; Gaps 1;
 QY 17 MWLLILVAYMFO-RNVNSVHMPKAVDPEAFNISEIIHQGYPCPEYEVATEDGYILSV 75
 Db 1 MWLLILVAYMFO-RNVNSVHMPKAVDPEAFNISEIIHQGYPCPEYEVATEDGYILSV 60
 QY 76 NRIIPRGLVQPKTGSRPVVLLQGLVGGASWISNLPNNISLGFILADAGFDVWGNISRGNA 135
 Db 61 YRPHGKNNSNIGKPEPVVYLQGLVGGASWISNLPNNISLGFILADAGFDVWGNISRGNA 120
 QY 136 AWRKHHTLSIDODEFWAFSYDEMARFDLPAVINFILOKQGOEKIYVGVYSGQTTWGFIA 195

Db 121 TWSRKNVYSPDSVEFWAFSDEWAKYDLPATINFIQKTGOEKIHYVGHVSQRTTIGFIA 180
 QY 196 FSTWPELAQIKMYFALAPIATVHKASPGTKPFLLLPDMNKGFLGKKEFLYOTRFLRQ 255
 Db 181 FSTWPELAQIKMYFALAPIATVHKASPGTKPFLLLPDMNKGFLGKKEFLYOTRFLRQ 240
 QY 256 VIYLCGOVILDOICSNIMLLGGFNTNNMMSASVYAAHTLAGTSVONILHWSQAVNSG 315
 Db 241 GTEVCSREVLDDLCSTNLFIFCGFDKKNLNVSRPDVYLGHNPAGTSVQDFLHWAQLVRSG 300
 QY 316 ELRAFDWGSSTKMLEKCNQPTPVRYVRDNTVPTAMWTGGQDLWLSNPEDVYKMLLSEVTNL 375
 Db 301 KFOAFNMGSPSQNNLHYNQKTPPEYDVSAMTVPAVWNGGNDILADPQDVAMLLPKLSNL 360
 QY 376 IYHKNIPEWAHVDPIWGLDAPHMYNEIHLMOOE 410
 Db 361 LFHKEILAYNHLDFIWANDAPQEVYNEIMISMAED 395

Search completed: February 19, 2004, 14:20:36
 Job time : 37.8598 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2004, 14:17:49 ; Search time 14.5693 Seconds
(without alignments)
2792.326 Million cell updates/sec

Title: US-10-042-431-47

Perfect score: 2247

Sequence: 1 MLETLRSQWIVSHRMENWLL.....IHLMQQETNLSQRCRAVL 423

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Pir1.*

2: Pir2.*

3: Pir3.*

4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1290	57.4	399	2	S41408 lysosomal acid lip
2	1290	57.4	399	2	G01416 lysosomal acid lip
3	1164.5	51.8	395	1	L1RTT triacylglycerol li
4	1139	50.7	397	1	JC4017 triacylglycerol li
5	1119	49.8	398	2	S07145 triacylglycerol li
6	752.5	33.5	403	2	T33198 hypothetical prote
7	751	33.4	411	2	T32290 hypothetical prote
8	730	32.5	405	2	T22675 hypothetical prote
9	707.5	31.5	405	2	H88930 protein R11G11.14
10	692	30.8	411	2	G89074 protein K04A8.5 [i
11	655.5	29.2	426	2	T20480 hypothetical prote
12	574.5	25.6	559	2	JT0949 egg-specific prote
13	505	22.5	443	2	T39540 triglyceride lipas
14	478.5	21.3	1585	2	T31611 hypothetical prote
15	475	21.1	467	2	T41053 triglyceride lipas
16	445	19.8	344	2	E84526 probable lysosomal
17	422.5	18.8	548	2	S37969 probable triacygl
18	412.5	18.4	460	2	T39443 probable triglycer
19	411.5	18.3	413	2	T43170 probable triacygl
20	294	13.1	431	2	D86318 protein P15H18.6 [
21	285	12.7	538	2	S84842 probable membrane
22	271.5	12.1	573	2	S84754 probable membrane
23	234.5	10.4	509	2	G36766 protein lipase F2P
24	149	6.6	62	2	S59904 lipase, pregastric
25	143	6.4	336	2	F83425 probable esterase/
26	126	5.6	987	2	A64474 hypothetical prote
27	113	5.0	412	2	B72391 conserved hypotet
28	112.5	5.0	40	2	S19533 triacylglycerol li
29	107	4.8	281	2	A11912 lysophospholipase

RESULT 1

S41408

Lysosomal acid lipase (EC 3.1.1.-) / sterol esterase (EC 3.1.1.13) precursor - human
C;Species: Homo sapiens (man)

C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 18-Jun-1999

C;Accession: S41408; A39315; S47187

R;Ameis, D.; Merkel, M.; Eckerskorn, C.; Greten, H.

Eur. J. Biochem. 219, 905-914, 1994

A;Title: Purification, characterization and molecular cloning of human hepatic lysosomal
A;Reference number: S41408; MUID:94155897; PMID:8112342

A;Accession: S41408

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-399 <AME>

A;Cross-references: EMBL:X76488; NID:G434305; PIDN:CAA54026.1; PID:G434306

R;Anderson, R.A.; Sando, G.N.

J. Biol. Chem. 266, 22479-22484, 1991

A;Title: Cloning and expression of cDNA encoding human lysosomal acid lipase/cholesteryl

A;Reference number: A39315; MUID:92042192; PMID:1718995

A;Accession: A39315

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-15, 'P', 17-399 <AND>

A;Cross-references: GB:M74775; NID:G187151; PIDN:AAA59519.1; PID:G187152

R;Du, H.; Gregory, G.A.

submitted to the EMBL Data Library, April 1994

A;Description: Structural conservation of putative functional motifs between mouse and h

A;Reference number: S47187

A;Accession: S47187

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-22, 'R', 24-399 <DUH>

A;Cross-references: EMBL:Z31690; NID:G506430; PIDN:CAA83495.1; PID:G506431

C;Genetics:

A;Gene: GDB:LIPA

A;Cross-references: GDB:120153; OMIM:278000

A;Map position: 10q24-10q25

C;Superfamily: triacylglycerol lipase, lingual

C;Keywords: carboxylic ester hydrolase; Glycoprotein

Query Match 57.4%; Score 1290; DB 2; Length 399;

Best Local Similarity 61.0%; Pred. No. 1.4e-105;

Matches 242; Conservative 59; Mismatches 94; Indels 2; Gaps 1;

QY 15 MEMWLLILVAYFQFNVSVMPTK--AVDPFANVISEIITOHQGPCEEEVATEDGYI 72

DB 1 MKMRFLGLVVCILVTLHSEGGGKLTAVDPETNNVSEIISYWGFPSEEVLTEDGYI 60

QY 73 LSVNRIPGLVQPKTKGRPVVLLQHLGVGGASNWSLNPNNLSGLFILADAGFDVWNGS 132

DB 61 LCLNRPKRNKNSDKGPKPVVFLQHLGLADSSNNVTNLANSGLFILADAGFDVWNGS 120

QY 133 RGNWSRKHHTLSIDODEFWAFSDYDEWARFDLPVAVINFILOKTOGQEKIYVYGVSGQTWVG 192
 Db 121 RGNWSRKHHTLSVSDODEFWAFSDYDEWAKYDLPASINFILNKTKGQGVYVGHSGQTITG 180
 QY 193 FIAFTWPELAQIKMYFALAPATATVKAHSPGTEKLLLPDMWIKGLFGKKEFLYQTRFL 252
 Db 181 FIAFSQIPELAKRIKMFALGPVAVAFCTSPWAKLGRPLDHLIKDLFGDKEFLPQSAFL 240
 QY 253 RQVIVLCGOVILDQICSNMILLGGFNTNMMNSRASVYAHTLAGTSVQNLHWSQAV 312
 Db 241 KWLGTHTVCTHVLKELCGNLCFLLCGFNERNLMSRDVYTHSPAGTSVQNLHWSQAV 300
 QY 313 NSGELRAFDSGSETKLEKCNQPTPVRYRVRTMTVPTAMWTGGQDMLSNPDEVMKLLSEV 372
 Db 301 KFQFQAFDGGSSAKNYFHYNQSPYPTNYKMLVPTAVWSGGHDLADYVDNILLTQI 360
 QY 373 TNLHYKNIPEWAHVDPIWGLDAPHRMYNEIHLMOQ 409
 Db 361 TNLVPHESIPWEHLDPIWGLDAPWRLYNKIINLMRK 397

RESULT 2

G01416
 lysosomal acid lipase - human
 C;Species: Homo sapiens (man)
 C;Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 18-Jun-1999
 C;Accession: G01416
 R;Du, H.
 submitted to the EMBL Data Library, April 1994
 A;Reference number: G06919
 A;Accession: G01416
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-399 <DUX>
 A;Cross-references: EMBL:U08464; NID:G505052; PIDN:AA60328.1; PID:G505053
 C;Superfamily: triacylglycerol lipase, lingual
 C;Keywords: Glycoprotein

Query Match 57.4%; Score 1290; DB 2; Length 399;
 Best Local Similarity 61.0%; Pred. No. 1.4e-105;
 Matches 242; Conservative 59; Mismatches 94; Indels 2; Gaps 1;
 QY 15 MEWMLLTVAYMFQNNVSNVHPTK--AVDPEAFPMNISEIIHQGYECREYEVATEDGYI 72
 Db 1 MKRFLGLVCLVLTWLSHSGSRGKLTAVDPETNMNVSEIISYWGPPSEYLVETEDGYI 60
 QY 73 LSVNRIPRGLVQPKTKTSRFPVLLQHLVGASNWSINLNNLSIGFILADAGFDVWVWGS 132
 Db 61 LCLNRI PHGRKNHSDKGPKEVWFLQHLGLADSSNWNVTNLANSLSIGFILADAGFDVWVWGS 120
 QY 133 RGNWSRKHHTLSIDODEFWAFSDYDEWARFDLPVAVINFILOKTOGQEKIYVYGVSGQTWVG 192
 Db 121 RGNWSRKHHTLSVSDODEFWAFSDYDEWAKYDLPASINFILNKTKGQGVYVGHSGQTITG 180
 QY 193 FIAFTWPELAQIKMYFALAPATATVKAHSPGTEKLLLPDMWIKGLFGKKEFLYQTRFL 252
 Db 181 FIAFSQIPELAKRIKMFALGPVAVAFCTSPWAKLGRPLDHLIKDLFGDKEFLPQSAFL 240
 QY 253 RQVIVLCGOVILDQICSNMILLGGFNTNMMNSRASVYAHTLAGTSVQNLHWSQAV 312
 Db 241 KWLGTHTVCTHVLKELCGNLCFLLCGFNERNLMSRDVYTHSPAGTSVQNLHWSQAV 300
 QY 313 NSGELRAFDSGSETKLEKCNQPTPVRYRVRTMTVPTAMWTGGQDMLSNPDEVMKLLSEV 372
 Db 301 KFQFQAFDGGSSAKNYFHYNQSPYPTNYKMLVPTAVWSGGHDLADYVDNILLTQI 360
 QY 373 TNLHYKNIPEWAHVDPIWGLDAPHRMYNEIHLMOQ 409
 Db 361 TNLVPHESIPWEHLDPIWGLDAPWRLYNKIINLMRK 397

RESULT 3

LIRTT

triacylglycerol lipase (EC 3.1.1.3) precursor, lingual - rat
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 11-Jun-1999
 C;Accession: A23045
 R;Doeherty, A.J.P.; Bodmer, M.W.; Angal, S.; Verger, R.; Riviere, C.; Lowe, P.A.; Lyons, Nucleic Acids Res. 13, 1891-1903, 1985
 A;Title: Molecular cloning and nucleotide sequence of rat lingual lipase cDNA.
 A;Reference number: A23045; MUID:85215587; PMID:3839077
 A;Accession: A23045
 A;Molecule type: mRNA
 A;Residues: 1-395 <DOC>
 A;Cross-references: GB:X02309; NID:G56595; PIDN:CRA26179.1; PID:G56596
 A;Experimental source: strain Sprague-Dawley
 A;Note: The partial sequence of the mature protein from a different, unspecified strain
 C;Comment: This acid-stable lipase is secreted by the serous (von Ebner's) glands at the
 C;Superfamily: triacylglycerol lipase, lingual
 C;Keywords: carboxylic ester hydrolase; glycoprotein; lipid digestion; saliva; serous gl
 F11-18/Domain: signal sequence #status predicted <SIG>
 F19-395/Product: triacylglycerol lipase, lingual #status predicted <MPT>
 F133,68,98,184,270/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 51.8%; Score 1164.5; DB 1; Length 395;
 Best Local Similarity 54.9%; Pred. No. 1.5e-94;
 Matches 217; Conservative 69; Mismatches 108; Indels 1; Gaps 1;
 QY 17 MWLLILVAYMFQ--RNVNSVEMETKAVDPEAFPMNISEIIHQGYECREYEVATEDGYILSV 75
 Db 1 MWLLILTVISTFGGAHGLFGKLGPGNPEANMISOMITWYVPCQYEVVTTEDGYILGV 60
 QY 76 NRIPRGLVQPKTKTSRFPVLLQHLVGASNWSINLNNLSIGFILADAGFDVWVWGSRG 135
 Db 61 YRIPHGKNNSENIGKRPVYVLOHGLIASATNWTANLNNLSIGFILADAGFDVWVWGSRG 120
 QY 136 AWSRKHHTLSIDODEFWAFSDYDEWARFDLPVAVINFILOKTOGQEKIYVYGVSGQTWGFIA 195
 Db 121 TWSRKNVYSPDSVEFWAFSEDEWAKYDLPATINFILOKTOGQEKIYVYGVSGQTWGFIA 180
 QY 196 FSTWPELAQIKMYFALAPATATVKAHSPGTEKLLLPDMWIKGLFGKKEFLYQTRFLRQL 255
 Db 181 FSTWPTLAKIKTYALAPVATVKTQSPKLSIFPTFLKLMFGKKEFLPHTYDFDPL 240
 QY 256 VYLICGVILDOICSNMILLGGFNTNMMNSRASVYAHTLAGTSVQNLHWSQAVNSG 315
 Db 241 GTEVCSREVLDDLSNLTIFFCGDFDKNLNVSFDPVILGNPAGTSVQDLFLHQAQLVRS 300
 QY 316 ELRAFDMGSETKLEKCNQPTPVRYRVRTMTVPTAMWTGGQDMLSNPDEVMKLLSEVTL 375
 Db 301 KFOAFNWSGSPQNMLHYNQKTPPEYDVSAMTPVAVWVWGGNDILADPDQVAMLLPKLSNL 360
 QY 376 LYHKNIPWAHVDPIWGLDAPHRMYNEIHLMOQ 410
 Db 361 LFKHEILAYNHLDFIAMDAPQEVINEMISMAED 395

RESULT 4

JC4017
 triacylglycerol lipase (EC 3.1.1.3) PGE precursor - bovine
 N;Alternate names: pregastric esterase
 C;Species: Bos primigenius taurus (cattle)
 C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C;Accession: JC4017; S64678
 R;Timmermans, M.Y.J.; Teuchy, H.; Kupers, L.P.M.
 Gene 147, 259-262, 1994
 A;Title: The cDNA sequence encoding bovine pregastric esterase.
 A;Reference number: JC4017; MUID:95011625; PMID:7926811
 A;Accession: JC4017
 A;Molecule type: mRNA
 A;Residues: 1-397 <TIM>
 A;Cross-references: GB:L26319; NID:G600756; PIDN:AAA57037.1; PID:G600757
 A;Experimental source: tongue
 R;Timmermans, M.Y.J.; Reekmans, G.; Teuchy, H.J.H.; Kupers, L.P.M.
 Biochem. J. 314, 931-936, 1996
 A;Title: Inhibition studies on calf pregastric esterase: the enzyme has no functional t

A;Reference number: S64678; MUID:96177869; PMID:8615791

A;Accession: S64678

A;Molecule type: protein

A;Residues: 24-40,248-253 <ITW>

C;Comment: Pregastric esterase is a major fat-digesting enzyme.

C;Genetics:

A;Gene: pge

C;Superfamily: triacylglycerol lipase, lingual

C;Keywords: blocked amino end; carboxylic ester hydrolase; glycoprotein

F;1-19/Domain: signal sequence #status predicted <SG>

F;20-397/Product: pregastric esterase #status predicted <MAT>

F;33,270,326/Binding site: carbohydrate (Asn) (covalent) #status predicted

F;171/Active site: Ser #status predicted

Query Match 50.7%; Score 1139; DB 1; Length 397;
Best Local Similarity 53.7%; Pred. No. 2,7e-92;
Matches 212; Conservative 68; Mismatches 113; Indels 2; Gaps 2;

QY 18 WLLILVAYMFORNNSVHPTKAVDPEAFNISEIIHQHGPCPEEYEVATEDGVLISVNR 77

DB 3 WLLVTVCFIHNHSGNAFCFLGIAKNPEASNMVSMISYGFSEMHKVTADGVLQVYR 62

QY 78 IPRGLVQPKTKGSRPVLLQHLGVGASNWIISNLPNNSLGLFLLADAGFDVWNGSRGNW 137

DB 63 IPHGKNANHLQORPVVFLQHLGLGSATNWIISNLPNNSLGLFLLADAGFDVWNGSRNTW 122

QY 138 SRKHKTLSIDODEFWAFSYDEMARFDLPVAVINFILOKTQGEKIYYVGSQTTMGFTAFS 197

DB 123 AQEHLVYSPDSPEFWAFSDFDEMAEYDLPSTDFILRRTGQKHLVHVSQTTIGFTAFS 182

QY 198 TMEPLAOKIKMYFALAPATVATYKHAKEPCTKELLAPDMWIKLFGCKEFLYQTRFQ 256

DB 193 TSPTLAEKIKVYALAPVATYKTSKLNKALIPHELKFIIFGDKMP-YPHFTFLEQFLG 241

QY 257 IYLCQGVILDTICSNIMLLGFGFTNNMNSRASVYAAHTLAGTSVQNLHWSQAVNSGE 316

DB 242 VEMCSRETLDVLCNALFAITGVNKNFNSRLDVIYLAHNPAGTSVQNTLHWRQAVKSGK 301

QY 317 LRAFPWGETKNEKCNQOPTPVVRVRDMVTPTAMWTGGQDLNPNEDVQMLSEVNTLI 376

DB 302 FQAPDWGAPYQNLHMYHQFTPIYNTLTAMNYPVIAVWSADNLLADPQDVDFLLSKLSNLI 361

QY 377 YKHNPEWAHVDFTWGLDPAHMYNEIHLMOQEE 411

DB 362 YHKEIPNYNHLDFIWMADAPQEVNIVSLMAEDK 396

RESULT 5

S07145

triacylglycerol lipase (EC 3.1.1.3) precursor, gastric - human

C;Species: Homo sapiens (man)

C;Date: 29-Jan-1993 #sequence revision 29-Jan-1993 #text_change 18-Jun-1999

C;Accession: S07145; S27102; S04942

R;Bodmer, M.W.; Antal, S.; Yarranton, G.T.; Harris, T.J.R.; Lyons, A.; King, D.J.; Pierd

Biochim. Biophys. Acta 909, 237-244, 1987

A;Title: Molecular cloning of a human gastric lipase and expression of the enzyme in yea

A;Reference number: S07145; MUID:87299724; PMID:3304425

A;Accession: S07145

A;Molecule type: mRNA

A;Residues: 1-398 <BODL>

A;Cross-references: EMBL:X05997; NID:G31771; PIDN:CAA29413.1; PID:G758063

A;Accession: S27102

A;Molecule type: protein

A;Residues: 20-24,'X',26,'X',28-33,'X',35-42,'X',44,'X',46-48,'X',50,'X',52,'X',54-56 <B

A;Note: it is uncertain whether Met-1 or Met-7 is the initiator

R;Bernbaeck, S.; Blaackberg, L.

Eur. J. Biochem. 182, 495-499, 1989

A;Title: Human gastric lipase. The N-terminal tetrapeptide is essential for lipid bindin

A;Reference number: S04942; MUID:89325292; PMID:2753032

A;Accession: S04942

A;Molecule type: protein

A;Residues: 20-45 <BEP>

C;Superfamily: triacylglycerol lipase, lingual

C;Keywords: carboxylic ester hydrolase; extracellular protein; glycoprotein; lipid digest

F;1-19/Domain: signal sequence #status predicted <SG>

F;20-358/Product: triacylglycerol lipase, gastric (covalent) #status predicted <MAT>

F;34,99,185,271/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 49.8%; Score 1119; DB 2; Length 398;

Best Local Similarity 52.2%; Pred. No. 1.6e-90;

Matches 210; Conservative 73; Mismatches 107; Indels 12; Gaps 2;

QY 17 WLLILVAYMFORNNSVHPTKAV-----DPEAFNISEIIHQHGPCPEEYEVATED 69

DB 1 WMLLLTMSLI-----SVLGTTHGLFGKLPGPSVPTWNIISQMITYWGYPNEEYEVATED 55

QY 70 GTILSVNRIIPRGLVQPKTKGSRPVLLQHLGVGASNWIISNLPNNSLGLFLLADAGFDVWM 129

DB 56 GYILEVNRIPYKKNKSGNTGQRPVFLQHLGLASATNWIISNLPNNSLGLFLLADAGFDVWL 115

QY 130 GNSRGNASRKHTLSIDODEFWAFSYDEMARFDLPVAVINFILOKTQGEKIYYVGSQGT 189

DB 116 GNSRGNWARRNLYTSPDSVEFWAFSDFDEMAKYDLPATIDFIVKTKGQKQLHYVGHSGQT 175

QY 190 TMGFTAFSTMPLEAOKIKMYFALAPATVATYKHAKEPCTKELLAPDMWIKLFGCKEFLYQT 249

DB 176 TIGFTAFSTNPSLAKRIKTFYALAPVATYKTSKLNKLPVFPQSLFKFPGDKLIFYPHN 235

QY 250 RFLRQVILYLCQGVILDTICSNIMLLGFGFTNNMNSRASVYAAHTLAGTSVQNLHWS 309

DB 236 FFDQFLATEVCSREMLNLLCSNALFIICGFSKNTSRLDVLVLSHNPAGTSVQNNFWT 295

QY 310 QAVNSGELAFDWSGETKNEKCNQOPTPVVRVRDMVTPTAMWTGGQDLNPNEDVQML 369

DB 296 QAVKSGFQAYDWSGSPVQNMHYDQSPYPYNTVNTVPIAVWNGKOLLADPQDVGLLL 355

QY 370 SEVTLYIHKNPEWAHVDFTWGLDPAHMYNEIHLMOQEE 411

DB 356 PKLPNLIYHKEIPNYNHLDFIWMADAPQEVNIVSMISEDK 397

RESULT 6

T33198

hypothetical protein ZK6.7 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000

C;Accession: T33198

R;Wu, X.

submitted to the EMBL Data Library, May 1998

A;Description: The sequence of C. elegans cosmid ZK6.

A;Reference number: Z21301

A;Accession: T33198

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-403 <WUX>

A;Cross-references: EMBL:AF067942; PIDN:AAAC17694.1; GSPDB:GN00023; CESP:ZK6.7

A;Experimental source: strain Bristol N2; clone ZK6

C;Genetics:

A;Gene: CESP:ZK6.7

A;Map position: 5

A;Intons: 29/3; 63/1; 219/3; 319/3; 365/3

C;Superfamily: triacylglycerol lipase, lingual

Query Match 33.5%; Score 752.5; DB 2; Length 403;

Best Local Similarity 41.1%; Pred. No. 3e-59;

Matches 168; Conservative 66; Mismatches 142; Indels 33; Gaps 12;

QY 17 MW-LLILVAYMFORNNSVHPTKAVDPEAFNISEIIHQHGPCPEEYEVATEDGYILSV 75

DB 1 MRPVAVFLAAFTVQDVVGSN-----GDPELHMTTPTQIIRWGPYPMIYVATDDGYILEM 55

QY 76 NRIPRGLVQPKTK-----GSRPVLLQHLGVGASNWIISNLPNNSLGLFLLADAGFDVWM 129

DB 56 HRIIPG-----KTNVTPWNGKEPVPVFMQHLGCSSDWMVNLDPQSAGFLFADAGFDVWL 110

QY 130 GNSRGNASRKHTLSIDODEFWAFSYDEMARFDLPVAVINFILOKTQGEKIYYVGSQGT 189

Db 111 GNRGNTYNNKHKLPKSPSHAFWDSWDEMATYDNLAMINHVLEVTGQDSVYMGHSQOT 170
 QY 190 TMGFIATSTWP-ELAQIKIKYFALAPATVYKHAQSPGKFLLLPDMMIKG---LFGKKEP 245
 Db 171 LTMFSLKSDGDSFAKKIKKFPALAPIGSVKHKIGFLSPFANYFSLFDFGWFDFIAGGEF 230
 QY 246 LYQTRFLRQLVYLCCGVILD-QICSNIMLLGGFNTNNMNSRASVVAHAHTLAGTSVON 304
 Db 231 LNNWAKMAKADICGLKVEADLDCNVFLLAGPESDQWQTRVPVATHDPAQTSTQN 290
 QY 305 ILHWSQAVNSGELRAFDPWGSSEKLEKCNQPTPVRYRVRDMT---VFTAMWTGGQDMLS 360
 Db 291 IVEHWQMVHGGVPAYDWGKT-KKKYGOANPPEY---DFTAIGTKIYLYWSADMLA 346
 QY 361 NPEDEV-KMLLSEVNTLIYHKN--IPWAHVDFIWLGLDAPHRMYNEIHL 406
 Db 347 DTPDVPDYLLTRLNPALVAQNHLDPDYNHLDFTWGLRAPDDIYRPAIKL 395

RESULT 7

T22290
 hypothetical protein F46B6.8 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
 C:Accession: T22290
 R:McMurray, A.
 submitted to the EMBL Data Library, April 1996
 A:Reference number: Z19542
 A:Accession: T22290
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-411 <MIL>
 A:Cross-references: EMBL:Z70780; PIDN:CAA94824.1; GSPDB:GN00023; CESP:F46B6.8
 A:Experimental source: clone F46B6
 C:Genetics:
 A:Gene: CESP:F46B6.8
 A:Map position: 5
 A:Introns: 35/3; 69/1; 104/1; 226/3
 C:Superfamily: triacylglycerol lipase, lingual

Query Match 33.4%; Score 751; DB 2; Length 411;
 Best Local Similarity 40.4%; Pred. No. 4.2e-58;
 Matches 163; Conservative 72; Mismatches 138; Indels 30; Gaps 12;
 QY 20 LILVAYFQNVNSVHMTKAVDPEAFNISEIIHQGYPCPEYEYVATEDGYILS 79
 Db 10 LICTSFLLQITINA-HPDD---DPELNMNTSQIIRWGYKAEVHTVTTEDGYILOMCRIP 65
 QY 80 RGLVQPKKT-----GSRPVLLQHLVGGASNLSNLPNNSLGFILADAGFDVWNGSR 133
 Db 66 YG-----KTSVTPWNGKRPVILLQHLGLLACASDVNLLPTQSAFVYFADAGFDVWLGVR 120
 QY 134 GNWSRKHKTLSDQDWFAPSYDEMARFPLPAVINFILOKTKQEKIYVYVSGTGMGF 193
 Db 121 GTTYGRKNTKLDPSSETAFQFSWDEMAQYDVPAMVDHVLAMTQENLYTMGHSQGTLMF 180
 QY 194 --TAFSTWPELAQIKMYFALAPATVYKHAQSPGKFLLLPDMMIKG---LFGKKEFLYQ 248
 Db 181 THLAKOTDGGSFACKIKKFPALAPIGAVNKIGFLSPFAHKPSPEFGWTFELSGKDFLDP 240
 QY 249 TRFLRQLVYLCCGVILD-QICSNIMLLGGFNTNNMNSRASVVAHAHTLAGTSVONILH 307
 Db 241 NWTMAAKADICGASKEAEELCNELFLIGPESDQWNASRTAIYSSQDPAGTSTQNIHV 300
 QY 308 WSQAVNSGELRAFDPWGSSEKLEKCNQPTPVRY---RVDMTVPYATMTGQDMLSNPDP 364
 Db 301 WMQVNRGRVPAPDWGKKI-KKKYQGDTPPEYDFGAIKGTKI--HLYWSDDDWGLDGDPTD 357
 QY 365 V-KMLLSEVNTLIYHKN--PEWAHVDFIWLGLDAPHRMYNEII 404
 Db 358 IHDPELLAKELAPAVIAENVNLEKYNHLDFTWGLNATFQIYDKAI 400

RESULT 8

T22675
 hypothetical protein F54F3.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
 C:Accession: T22675
 R:Percy, C.; Lloyd, C.
 submitted to the EMBL Data Library, September 1996
 A:Reference number: Z19598
 A:Accession: T22675
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-405 <MIL>
 A:Cross-references: EMBL:Z79696; PIDN:CAB01973.1; GSPDB:GN00023; CESP:F54F3.3
 A:Experimental source: clone F54F3
 C:Genetics:
 A:Gene: CESP:F54F3.3
 A:Map position: 5
 A:Introns: 31/3; 65/1; 274/3; 366/3
 C:Superfamily: triacylglycerol lipase, lingual

Query Match 32.5%; Score 730; DB 2; Length 405;
 Best Local Similarity 38.6%; Pred. No. 2.9e-56;
 Matches 161; Conservative 69; Mismatches 155; Indels 32; Gaps 11;
 QY 15 MEMMLLILVAYFQNVNSVHMTKAVDPEAFNISEIIHQGYPCPEYEYVATEDGYILS 74
 Db 1 MRSWSTVMLAVLATAATVFGH---DADPEMKMTTPQIIMRWGYPAMIYDVTTEDGYILE 56
 QY 75 VNRTPRGLVQPKKT-----GSRPVLLQHLVGGASNLSNLPNNSLGFILADAGFDVW 128
 Db 57 LHRIPYG-----KTNVTPWNGKRPVFMQGLSCSSNNVVNLTPTSAFLPADAGYDWN 111
 QY 129 MGNRGNNAWRKHKTLSDQDWFAPSYDEMARFPLPAVINFILOKTKQEKIYVYVSGQ 188
 Db 112 LGNPRGNTYSKMKHKNLKPSSHAFWDSWDEMQQYDLPAMIEKALEVTGQDSLVIYHGSQ 171
 QY 189 T-TWGFIAFSTWPE---LAQIKMYFALAPATVYKHAQSPGKFLLLPDMMIKG---LF 240
 Db 172 TLTM---FSRLSEDKVGVGNKIKKFPALAPVGSVXHKIKALKFPADYSLFDFGWFDFV 227
 QY 241 GKKEFLYQTRFLRQLVYLCCGVILDQICSNIMLLGGFNTNNMNSRASVVAHAHTLAG 299
 Db 228 GSGEFLNNIMKLVSSEVCAGLKVERAGVDDVWELIAGPESQNLNATRPVIVAHTPAG 287
 QY 300 TSVONILHWSQAVNSGELRAFDPWGSSEKLEKCNQPTPVRYRVRDMTVPYATMTGQDML 359
 Db 288 TSTQNIYHWTQMVNRHGGTPKYDYG-EKGNKKHYGOANVPAYDFTTVNRFPVLYWGSDDL 346
 QY 360 SNPEDEV-KMLLSEV--TNLIYHKNIPWAHVDFIWLGLDAPHRMYNEIHLMOQEBTN 413
 Db 347 ADPTDVTDFLLTHLNPSTVVQNNKLDYNNHLDFIWLGLRAPKDIYEPIDIVRNDVLN 403

RESULT 9

H88930
 protein R11G11.14 [imported] - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
 C:Accession: H88930
 R:Anonymous, The C. elegans Sequencing Consortium.
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
 A:Reference number: A75000; MUID:99069613; PMID:9851916
 A:Note: see websites: genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/projects/C.elegans/
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
 A:Accession: H88930
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-405 <STO>
 A:Cross-references: GB:chr V; PIDN:AAC69088.1; PID:G2384863; GSPDB:GN00023; CESP:R11G11.14
 A:Note: Similar to lipase; R11G11.14

C;Genetics:
A;Gene: Rllg1.14
A;Map position: 5
C;Superfamily: triacylglycerol lipase, lingual

Query Match 31.5%; Score 707.5; DB 2; Length 405;
Best Local Similarity 39.0%; Pred.No.2.8e-54;
Matches 162; Conservative 75; Mismatches 133; Indels 45; Gaps 15;

Qy 19 LLILVYMFORNYSNVMPTKAVDPPEAFNMISIIHQGYPCPEYEYATEDGYLSVNRI 78
|:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db 9 LVLILA-----VENHAKS---DPELHMTPOIIERWGWGPAMIYSTTDDGYLELHI 59
|:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Qy 79 PRGLVPQPKTI-----GSRPVLLLOGLVGASNWTSNLNNSLGTILADAGFDVMGNS 132
|:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db 60 PHG-----KTNVTWPNGKQPWFVWGGLLCASDTWTNDLPFQSAAFIFADAGFDVWLGNM 114
|:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Qy 133 RGNWSRKHKTLSIDODEFWAFPSYDEMAREPDLPAVINFILOKTOGOEKIYYGVYSGQT-TM 191
|:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db 115 RGNTYSMKHNKLASHSDFWESWDEMATYDLLPAMIDKULEVTIGQSSLYVMGHSQGTLTM 174
|:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Qy 192 GFATFSMP-----LAQIKMYPALAPITVKIAKSPGYKFLLLPDMMKMG----LFKKKE 244
|:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db 175 -----FSELSKDDGIFAKKIKFPALAPVGSVDKIKGFLSFPAHYPSLEDFDGFVEFCAGE 230
|:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Qy 245 FLYOTRFRLQIVLYCGOV--ILDQICSNIMLLGGFTNNMNMSRASVVAAHTLACTSVQ 303
|:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db 231 FLPNWAMKLAADI CGGLKIESDLDNVCFIAGPESDOWNSTRVPVYASHDPAGTATQ 290
|:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Qy 304 NILHWSQVNSBELRAPDWGSETKNLEKC-N-QPTPVRYRVDMT-----VPTAMTGGQDM 358
|:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db 291 NIWHVIQMWRHGVPAYDWGSK-ENKKNVFNQANPPFY---DFTAIGTKQIVLYWSDADW 346
|:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Qy 359 LSNPEDV-KMLSELVNLIIYHK-N--IPWAHVDFIWGLDAPHRMVNEITHMQOE 410
|:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db 347 LADKTDITNYLLTRNPAILTAQNFFTYDNHFDFVGLRAPNDIYLPIDVICTKD 401
|:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|

RESULT 10
G89074
protein K04A8.5 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
C;Accession: G89074
R;anonymous, The C. elegans Sequencing Consortium.
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biological processes
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_elegans/
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and Science 283, 2103, 1999;
A;Accession: G89074
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-411 <STO>
A;Cross-references: GB:chr_V; PIDN:AAC48051.1; PID:g1658354; GSPDB:GN00023; CESP:K04A8.5
C;Genetics:
A;Gene: K04A8.5
A;Map position: 5
C;Superfamily: triacylglycerol lipase, lingual

Query Match 30.8%; Score 692; DB 2; Length 411;
Best Local Similarity 36.8%; Pred.No.6.6e-53;
Matches 153; Conservative 83; Mismatches 154; Indels 26; Gaps 10;

Qy 11 VSHRMEWMILLILVAYMFQRNVNSVHMTPKAVIDPEAFNISEIHHOGYPCEEYEVATEDG 70
|:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db 5 ISDLMTVMPLLIILLLSNY-----SKSVDFEYLDTPELIKSWGYSVEIYNTTTKDG 57
|:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Qy 71 YILSVNEIRPGLVQPKT-----GSRPVLLLOGLVGASNWISNLPNNSLGFIADAGFDV 127
|:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db 58 FILELHRIPIGREYFTSSDYNNSRRVFILOHGLFCSSFDFVANSFHQSAGFVADAGFDV 117
|:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Qy 128 WMGNSRGNAWSRKHTLSIDOD-EFWAFPSYDEMARFDLPVAINFILOKTOGOEKIYYGVYGS 186
|:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|

RESULT 12

JT0949
egg-specific protein - silkworm
C:Species: Bombyx mori (silkworm)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 26-Feb-1998
C:Accession: JT0949; A28527
R:Sato, Y.; Yamashita, O.
submitted to JIPID, September 1991
A:Reference number: JT0949
A:Accession: JT0949
A:Molecule type: DNA
A:Residues: 1-559 <SAT>
A:Experimental source: larva
A>Note: this protein is a homotrimer
R:Indrasith, L.S.; Sasaki, T.; Yamashita, O.
J. Biol. Chem. 263, 1045-1051, 1988
A:Title: A unique protease responsible for selective degradation of a yolk protein in B.
A:Reference number: A28527; MUID:88087166; PMID:3275655
A:Accession: A28527
A:Molecule type: protein
A:Residues: 13-40;133-144, 'T', 146-152;229-248 <IND>
C:Genetics:
A:Map position: 19
C:Keywords: egg yolk; homotrimer
F:132-133/Cleavage site: Lys-Asn (egg 30.5K cysteine proteinase) #status experimental
F:228-229/Cleavage site: Arg-Asp (egg 30.5K cysteine proteinase) #status experimental

Query Match 25.6%; Score 574.5; DB 2; Length 559;
Best Local Similarity 33.2%; Pred. No. 2.3e-42;
Matches 132; Conservative 82; Mismatches 152; Indels 31; Gaps 9;

QY 32 NSVH---MPTKAVDPEAF-WNISEIIOHQGYPEEVEVATEDGYIISVNRIPRGLVQPKK 87
DB 171 NTFDAISETQRENEDFLNATELLKKHQPYVEEHTVATDDGYHITVLRIIP-PTHQTRD 229

QY 88 TGSPPVLLQHLGVGGASNISLNPNNLSGLFADAGFDVWNGNSRGNASRKHKTLSID 147
DB 230 DKKFPVALLHGLLSADLLWLPKPSKSLAYMLSDAGYDVLGNVGNKYSRSHVSKHPA 289

QY 148 QDEFWAFSYDEMARFDLPVAVINFLQKTGQEKIYYVYSGQTTMGFIAPSTMPPELAQKIK 207
DB 290 LNDPWKFSNDEIALHDLPAIDHVDLISGQERLHYIGHSQCATTFFALMSEQPSYNEKIV 349

QY 208 MYFALAPIATVYKAKS-----PQTKFL-LLPDMTKLGFCKEFLYQTRFLRLQVLYL 259
DB 350 SMHSLSPVYVYVRSPLFRMIAPTSKPYOYIHQVGHGAFEPGKHLIET-----F 400

QY 260 CGQVILDQ-----ICSNIMLLGGFTNNNMNRASVYAAHTLAGTSVQVNIHWSQAVNS 314
DB 401 GGAACREKLGCRHVCNNLNVISGINVYNQDADIVPVVMAHLPAAGTSARVMKQYQGNVAS 460

QY 315 GELRAPFGWSETKLEKCNQPTPVYRVDMVTPTAMWTGGQDWLSPEDVYKMLSEVTN 374
DB 461 HDPKRYNGAET-NMKYVGASEPPSYDLKSVAPVNLVYHSHDADLAPKDKVEXLQENLPN 519

QY 375 LIYHKNIPENWAH---VDFIWGLDAPHRYMNEIHLMQ 408
DB 520 VKQSFVEPEQOHFTDLDQFSKAPDVIYQKLMENMQ 556

RESULT 13
T39540
triglyceride lipase-cholesterol esterase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: T39540
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Beck, A.; Reinhardt, R.
submitted to the EMBL Data Library, February 1998
A:Reference number: Z21862
A:Accession: T39540
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

A;Residues: 1-443 <WOO>

A;Cross-references: EMBL:AL021748; PIDN:CAAL16863.1; GSPDB:GNO0067; SPDB:SPBC16A3.12c

A;Experimental source: strain 972h-; cosmid c16A3

C;Genetics:

A;Gene: SPDB:SPBC16A3.12c

A;Map position: 2

A;Introns: 11/2; 152/3; 282/1; 398/3

C;Superfamily: triacylglycerol lipase, lingual

Query Match 22.5%; Score 505; DB 2; Length 443;

Best Local Similarity 29.6%; Pred. No. 2.1e-36;

Matches 125; Conservative 91; Mismatches 154; Indels 52; Gaps 13;

QY 9 WIVSHRMEMW--LLILVAYMFORNVNSVHMPTKAVDPEAFMNI-SBIIHQGYCEYEVA 66

DB 35 YMIPERVREWFVRVRIIVLYHYIYSSKTTDGMTDAV--QKCRNIYEICEAFGYRVEEHLVR 92

QY 67 TEDGYIISVNRIPRGLVQPKKT-GSRPVVLLQHLGVGGASNWIS-NLPNNLSGLFADAG 124

DB 93 TQDNFILLHRI-----THPKSQSHKREVYVYCHHGLMTNSLWVAVNESRSLFPVLIESG 148

QY 125 FDVWNGNSRGNASRKHKTLSIDQDEFWAFSYDEMARFDLPVAVINFLQKTGQEKIYYVG 184

DB 149 YDVLGNNGNKGYSRKHITYPKDEEFWNFSLDDMAFDPDITVDYILRETGREKLNYYIG 208

QY 185 YSQTTMGFIAPSTMPPELAQKIKMYFALAPIATVYKAKSPGTFELL--PDMIKLGLFGK 242

DB 209 FSGQTAAAMAALSINFLNDKNWIFIGLAPAYAPKGSFYVDYIVKVNPKIMYH-LFGR 267

QY 243 KEFLYQTRFLRLQVLYLGGVILDOICSNIMLLGGFTNNNMNRASVYAAHTLAGTSV 302

DB 268 RCLIPSVTFWQV----ICYPPIFVKIVDVSLKILFWDLNLSINQKLCGYAHLVSESSV 323

QY 303 QNHLWSQAVNSGELRAPD-----WGSE-----TNLEKCNQPTPVYRVDMVT 346

DB 324 KSVVFWLQIIKXCTFQLYDQDDMALLAGYGRHYQVLPFTNNI-KC----- 368

QY 347 VPTAMWTGGQDWLSPEDVYKMLSEVTNLYYHKNIPENWAHVDFWGLDAPHRYMNEIHL 406

DB 369 -PMILWGGQKTLINMEVMTALPFAKEV---SIAHYEHLDFLWQGVKEEVPVVIDA 424

QY 407 MQ 408

DB 425 LK 426

RESULT 14

T31611

hypothetical protein Y508A.g - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T31611

R:Steward, C.

submitted to the EMBL Data Library, September 1999

A:Reference number: Z21047

A:Accession: T31611

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A;Residues: 1-1585 <WIL>

A;Cross-references: EMBL:AL117200; NID:e1549770; PIDN:CAB55050.1; CBSP:Y50E8A.g

A;Experimental source: clone Y50E8A

C;Genetics:

A;Gene: CBSP:Y50E8A.g

A;Introns: 25/3; 60/1; 133/2; 217/3; 270/3; 337/2; 400/1; 746/2

Query Match 21.3%; Score 478.5; DB 2; Length 1585;

Best Local Similarity 30.6%; Pred. No. 2.8e-33;

Matches 118; Conservative 67; Mismatches 172; Indels 29; Gaps 9;

QY 44 EAFNNISEIIOHQGYCEYEVEVATEDGYIISVNRIPRGLVQPKKTG--SRPVVLLQHLV 101

DB 20 EETLDAADTISHGYTVYKHYVTDDGYTVQLQRIPVGRDRSILGSKPVPVFFMEGLF 79

QY 102 GGASNNISLNNLSGLFADAGFDVVMGNSGNVNSRKHKTLSIDQDEWAFSYDEMAR 161
 Db 80 GSSVHFLNLPQSAAYIFADAGFDVWLGIRGTETGLNHTSFSTGVNFMFNSLYEHSH 139
 QY 162 FDPVAVINFILOKTGQKYYVGYSGQTTMGFTAFSTMPLEL-----AQKIMYFALAPIAT 217
 Db 140 YDLRQIEYVLDVTRHESLFPVGHSGQTA---VMFARLAEADVWQSKIRVFPALGFTAG 196
 QY 218 VGHAKSPGTFKLLPLDMMIKGLFGKKEFLYQTRF-----LRQLVYLGGQVILDO 267
 Db 197 FMPLMP---FTLEENYLOALI---QFALDGRGILPVEIPRAIASKFAFCSSKFFTF 250
 QY 268 ICNIMILLGGFNT-NNMNSRASVVAHAHTLAGTSVQNIILHWSQAVNSGELRAFDMGSET 326
 Db 251 LCSAGPKVAAGIETLGVNDSPRIILSHPPSATSTLNLLHWMQIFKYHELRLDLGT-A 309
 QY 327 KMLEKNQPTFVRVYRVEDMTVPTAMWTGGQDWLSNPDVQMLLSEV--TNLIYHKNIPW 384
 Db 310 RNLIAVGQKQDAPRLIGNIIAQTILYPSKDDQITDEVDVREIIMQMGFGLIESYDLDF 369
 QY 385 ARVDFIWLGLDAPHRMNYNEIHLMOQE 410
 Db 370 THDFILGLRATDEVYKPIVYRIYKE 395

RESULT 15

T41053
 triglyceride lipase-cholesterol esterase - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2003
 C:Accession: T41053
 R: Hilbert, H.; Duesterhoeft, A.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, August 1998
 A:Reference number: Z21967
 A:Accession: T41053
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-467 <HLL>
 A:Cross-references: EMBL:AL031324; PIDN:CAA20447.1; GSPDB:GN00068; SPDB:SPCC1672.09
 A:Experimental source: strain 972h-; cosmid c1672
 C:Genetics:
 A:Gene: SPDB:SPCC1672.09
 A:Map position: 3
 A:Introns: 12/2; 164/3; 294/1
 C:Superfamily: triacylglycerol lipase, lingual

Query Match 21.1%; Score 475; DB 2; Length 467;
 Best Local Similarity 31.2%; Pred. No. 1e-33;
 Matches 125; Conservative 68; Mismatches 158; Indels 50; Gaps 12;
 QY 39 KAVDPEAFMISELIHQGYPCBEYEVATEDGYLSVNRIPR-----GLVQPKTGSRP 92
 Db 73 KAIHDA--DIRECKISGYVEDHLVTRDEDDYILCIHRISKDPSGRIGSPHPKKL---P 127
 QY 93 VLLQHLGLVGSAGNINL--PNNLSGLFADAGFDVVMGNSGNVNSRKHKTLSIDQDEF 151
 Db 128 VVYCHGLLNNSEVWCVNCPNCLVFDLVNKGVDVWLGNNRGNKYRCHLRFDSTDKEF 187
 QY 152 WAFSYDEMARFDLPVAVINFILOKTGQKYYVGYSGQTTMGFTAFSTMPLELAQIKYFA 211
 Db 188 WDFSIDDFQYDIPDTIDYILKTSQTKLYIGFSQTAQAFASLSIHPLLNDKINSLIA 247
 QY 212 LAPATVYKHAKSCTKELLPLDMMIKG-----LQCKEFLYQTRFLRQLVYLCQVT 264
 Db 248 LAP-----AISPGLHNRVVDVAFKARPSILFFLGRKSLPSAGFWQSFL----APKF 297
 QY 265 LDQICSNIMILLGGFNTNNMNSRASVVAHAHTLAGTSVQNIILHWSQAVNSGELRAFDP--- 321
 Db 298 FDAVLAYCLSQLFNWSQNISSYQRLVSYFAHLYSYTSVKLVHVFQIMRSABEPRMYDNDQ 357
 QY 322 WGSB--TKNLEKNQPTFVRVYRVEDMTVPTAMWTGGQDWLSNPDVQMLLSEVTNLIYHK 379
 Db 358 LGHDYFLKYKAAKFPT-----NNIRPTIYLIWGGSDSL---VDIQAMLNALPAEVEHV 408

QY 380 NIPWAHVDFWING-----LDAPHRMNYNEIHLMOQBETH 413
 Db 409 KVDSEHLDMTWADTVKDYVIPPVLRRLRDLHHPPEHEEND 449

Search completed: February 19, 2004, 14:23:06
 Job time : 19.5683 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2004, 14:16:19 : Search time 9.88561 Seconds
(without alignments)
2012.248 Million cell updates/sec

Title: US-10-042-431-47

Perfect score: 2247

Sequence: 1 MLETLRSQWIVSHRMENWLL.....IHLMQQETNLSQRCSEAVL 423

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1289	57.4	399	1 LICH_HUMAN	P38571 homo sapien
2	1220.5	54.3	397	1 LICH_RAT	Q84194 rattus norv
3	1197.5	53.3	397	1 LICH_MOUSE	Q520M5 mus musculu
4	1164.5	51.8	395	1 LIPG_RAT	P24634 rattus norv
5	1154	51.4	398	1 LIPG_CANFA	P80035 canis famil
6	1139	50.7	397	1 LIPG_BOVIN	Q29458 bos taurus
7	1119	49.8	398	1 LIPG_HUMAN	P07098 homo sapien
8	663	29.5	394	1 LIP3_DROME	O46108 drosophila
9	546	24.3	439	1 LIP1_DROME	O46107 drosophila
10	422.5	18.8	548	1 TGL1_YEAST	P34163 saccharomyc
11	126	5.6	987	1 YD94_MERTUA	Q58789 methanococc
12	105	4.7	430	1 ORC2_CAEEL	Q21037 caenorhabdi
13	101	4.5	436	1 PAPA_CAVPO	P70683 c platelet-
14	96.5	4.3	287	1 Y193_HAETN	Q57427 haemophilus
15	95.5	4.3	1259	1 L1N1_HUMAN	P08547 homo sapien
16	93.5	4.2	278	1 PRXC_STRAU	O31168 streptomyce
17	93	4.1	569	1 PRD_FLAPA	Q08210 plasmodium
18	92	4.1	748	1 PTIP_ECOLI	P37177 escherichia
19	92	4.1	748	1 PTIP_SALTY	P37178 salmonella
20	91	4.0	370	1 P2C2_SCHPO	Q09172 schizosacch
21	90.5	4.0	1037	1 ACRD_ECOLI	P24177 escherichia
22	90	4.0	277	1 BPA2_STRAU	P29715 streptomyce
23	90	4.0	1122	1 Y3C3_YEAST	P33280 saccharomyc
24	90	4.0	1136	1 POLG_OMV	Q9esd7 mus musculu
25	90	4.0	2083	1 DYSF_MOUSE	P33691 rhizobium m
26	89	4.0	330	1 EXOA_RHINE	P33691 rhizobium m
27	89	4.0	422	1 YG83_SHEON	P59352 shewanella
28	89	4.0	521	1 ACH1_SCHPO	Q9uni9 schizosacch
29	88.5	3.9	185	1 YCZH_BACSU	O31492 bacillus su
30	88.5	3.9	869	1 CFAC_ECOLI	P25733 escherichia
31	87	3.9	473	1 SCRB_LACIA	Q04937 lactococcus
32	86.5	3.8	382	1 YDHH_HAETN	P44861 haemophilus
33	86.5	3.8	530	1 SYN_ANASP	Q8z068 anabaena sp

RESULT 1

ID	SEQUENCE	STANDARD	PRT	399 AA
AC	P38571	Q16529; Q96EJ0;		
DT	01-OCT-1994	(Rel. 30, Created)		
DT	01-OCT-1994	(Rel. 30, Last sequence update)		
DT	15-SEP-2003	(Rel. 42, Last annotation update)		
DE	lysosomal acid lipase/cholesterol ester hydrolase precursor			
DE	(EC 3.1.1.13) (LAL) (Acid cholesteryl ester hydrolase) (Sterol			
DE	esterase) (Lipase A) (Cholesteryl esterase).			
GN	LIPA.			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RI	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 196-212; 277-297 AND 305-315.			
RX	MEDLINE=92042192; PubMed=1718995;			
RA	Anderson R.A., Sando G.N.;			
RT	"Cloning and expression of cDNA encoding human lysosomal acid			
RT	lipase/cholesterol ester hydrolase. Similarities to gastric and			
RT	lingual lipases.";			
RL	J. Biol. Chem. 266:22479-22484(1991).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver;			
RX	MEDLINE=94155897; PubMed=8112342;			
RA	Ameis D., Merkel M., Eckerskorn C., Greten H.;			
RT	"Purification, characterization and molecular cloning of human			
RT	hepatic lysosomal acid lipase.";			
RL	Eur. J. Biochem. 219:905-914(1994).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver;			
RX	MEDLINE=96363957; PubMed=8725147;			
RA	Du H., Witte D.P., Grabowski G.A.;			
RT	"Tissue and cellular specific expression of murine lysosomal acid			
RT	lipase mRNA and protein.";			
RL	J. Lipid Res. 37:937-949(1996).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Placenta;			
RX	MEDLINE=22388257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,			
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W.,			
RA	Villalón D.K., Muzny D.C., Harte S., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,			
RA	Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			

ALIGNMENTS

34	86.5	3.8	685	1	INVA_SALTY
35	86.5	3.8	1882	1	Y468_MYCPN
36	86	3.8	525	1	ACUB_NEUCR
37	86	3.8	1866	1	VGNB_CPMV
38	85.5	3.8	428	1	YF63_SCHPO
39	85	3.8	292	1	FABF_NEUCR
40	85	3.8	429	1	ARB2_ECOLI
41	85	3.8	429	1	BIOA_EUCBP
42	85	3.8	837	1	UPBA_DICDI
43	85	3.8	4128	1	PRKD_MOUSE
44	84.5	3.8	875	1	AMD2_XENLA
45	84	3.7	451	1	YB27_YEAST

P35657	salmonella
P75109	mycoplasma
P03600	neurospora
P03600	cowpea mose
O14249	schizosacch
Q9hgr3	neurospora
P52146	escherichia
Q89ak4	buchnera ap
P54201	dictyosteli
P97313	mus musculu
P12890	xenopus lae
P38295	saccharomyc

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [5]
 RP VARIANT CSD ARG-129 AND PRO-129.
 RX MEDLINE=94195814; PubMed=8146180;
 RA Anderson R.A., Byrum R.S., Coates P.M., Sando G.N.;
 RT "Mutations at the lysosomal acid cholesteryl ester hydrolase gene
 RT locus in Wolman disease.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:2718-2722(1994).
 RN [6]
 RP VARIANTS CSD ARG-129 AND PRO-129.
 RX MEDLINE=98295576; PubMed=9633819;
 RA Ries S., Buechler C., Schindler G., Aslanidis C., Amels D., Gasche C.,
 RA Jung N., Schambach A., Fehrer P., Vanier M.T., Belli D.C.,
 RA Greten H., Schmitz G.;
 RT "Different missense mutations in histidine-108 of lysosomal acid
 RT lipase cause cholesteryl ester storage disease in unrelated compound
 RT heterozygous and hemizygous individuals.";
 RL Hum. Mutat. 12:44-51(1998).
 CC -1- FUNCTION: CRUCIAL FOR THE INTRACELLULAR HYDROLYSIS OF CHOLESTERYL
 CC ESTERS AND TRIGLYCERIDES THAT HAVE BEEN INTERNALIZED VIA RECEPTOR-
 CC MEDIATED ENDOCYTOSIS OF LIPOPROTEIN PARTICLES. IMPORTANT IN
 CC MEDIATING THE EFFECT OF LDL (LOW DENSITY LIPOPROTEIN) UPTAKE ON
 CC SUPPRESSION OF HYDROXYMETHYLGLUTARYL-COA REDUCTASE AND ACTIVATION
 CC OF ENDOGENOUS CELLULAR CHOLESTERYL ESTER FORMATION.
 CC -1- CATALYTIC ACTIVITY: A sterol ester + H₂O = a sterol + a fatty
 CC acid.
 CC -1- SUBCELLULAR LOCATION: Lysosomal.
 CC -1- DISEASE: DEFECTS IN LIPA ARE THE CAUSE OF THE SEVERE INFANTILE-
 CC ONSET WOLMAN DISEASE (WD) AND THE Milder LATE-ONSET CHOLESTERYL
 CC ESTER STORAGE DISEASE (CESD).
 CC -1- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M74775; AAB59519.1; -;
 DR EMBL; U04285; AAB60327.1; -;
 DR EMBL; U04286; AAB60327.1; JOINED.
 DR EMBL; U04287; AAB60327.1; JOINED.
 DR EMBL; U04288; AAB60327.1; JOINED.
 DR EMBL; U04290; AAB60327.1; JOINED.
 DR EMBL; U04291; AAB60327.1; JOINED.
 DR EMBL; U04292; AAB60327.1; JOINED.
 DR EMBL; U04293; AAB60327.1; JOINED.
 DR EMBL; X76488; CAAS4026.1; -;
 DR EMBL; Z33690; CAAS4026.1; -;
 DR EMBL; U08464; AAB60328.1; -;
 DR EMBL; BC012287; AAH12287.1; -;
 DR PIR; S41408; S41408.
 DR Genew; HGNC:6617; LIPA.
 DR MIM; 278000; -;
 DR GO; GO:0005764; C:lysosome; TAS.
 DR GO; GO:0006487; P:N-linked glycosylation; TAS.
 DR InterPro; IPR000073; A/B hydrolase.
 DR InterPro; IPR000734; Lipase.
 DR InterPro; IPR000379; Ser. esters site.
 DR Pfam; PF04083; abhydro lipase; 1.
 DR Pfam; PF00561; abhydro lipase; 1.
 DR PROSITE; PS00120; LIPASE SER; 1.
 DR HydroLase; Lipid degradation; Glycoprotein; Signal; Lysosome;
 KW Disease mutation; Polymorphism.
 FT SIGNAL 1 21 POTENTIAL.

FT CHAIN 22 399 LYSOSOMAL ACID LIPASE/CHOLESTERYL ESTER
 FT ACT_SITE 174 174 HYDROLASE.
 FT ACT_SITE 374 374 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT CARBOHYD 36 36 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT CARBOHYD 72 72 N-LINKED (GLNAC. .) (POTENTIAL).
 FT CARBOHYD 101 101 N-LINKED (GLNAC. .) (POTENTIAL).
 FT CARBOHYD 161 161 N-LINKED (GLNAC. .) (POTENTIAL).
 FT CARBOHYD 273 273 N-LINKED (GLNAC. .) (POTENTIAL).
 FT CARBOHYD 321 321 N-LINKED (GLNAC. .) (POTENTIAL).
 FT VARIANT 16 16 P -> T.
 FT VARIANT 129 129 H -> P (IN CESD).
 FT VARIANT 129 129 H -> R (IN CESD).
 FT VARIANT 200 200 L -> P (IN CESD AND WD).
 FT CONFLICT 23 23 /FTID-VAR 004250.
 FT CONFLICT 29 29 G -> R (IN REF. 3).
 FT CONFLICT 29 29 V -> L (IN REF. 4).
 SQ SEQUENCE 399 AA; 45415 MW; 55F27391306B09A CRC64;
 Query Match 57.4%; Score 1289; DB 1; Length 399;
 Best Local Similarity 63.5%; Pred. No. 3.5e-103; Indels 0; Gaps 0;
 Matches 235; Conservative 55; Mismatches 80;
 QY 40 AVDPAPFNISEIIHQGYPCBEYEVATEDGYILSVNRIPIRGVLQPKTKGSRPVLQHG 99
 DB 28 AVDPETNMVSEIISYWGPPSEYLVETEDGYILCLNRIPIHGRKNHSDKGPKEVFLQHG 87
 QY 100 LVGASWISNLPNSLGFILADAGFDVWNGSRGNANSHKHTLSIDODEWAFSYDEM 159
 DB 88 LLADSSNWNVTLANSSILGFILADAGFDVWNGSRGNANSHKHTLSIDODEWAFSYDEM 147
 QY 160 ARPLPAVINFILOKQGEKIVYVYSGQTMGFIAFTWPELAQIKYFALAPIATVK 219
 DB 148 AKYDLPASINFILOKQGEKIVYVYSGQTMGFIAFTWPELAQIKYFALAPIATVK 207
 QY 220 HAKSPGTFKLLPDMTKLFGKKEFYIYQTRQLVYLCGOVILDDQICSNMILLGGF 279
 DB 208 PCTSPMAKRLPDLHKLDFGDKFELPQSAFLKWLGTHTVTHVILKELCGNLCFLCGF 267
 QY 280 NTNNMNSRASVYAAHTLAGTSVQNTLHWSOAVNSGELRAFDWGSSTKLEKCNQPTPVR 339
 DB 268 NERNLNRSVDVYTHSPAGTSVQNTLHWSOAVNSGELRAFDWGSSTKLEKCNQPTPVR 327
 QY 340 YRVRDVTPTAMWTGGQDNLSPEDVKLLSEVNTLIYHKNIPFWAHVDFWGLDAPHRM 399
 DB 328 YNVKMLVPTAVNSGCHDLADVVDVNNILLTQITNLVFEHSIPWEHLDFWGLDAPWRL 387
 QY 400 YNEIHLMOQ 409
 DB 388 YNKIINLMRK 397
 RESULT 2
 ID LICH RAT STANDARD; PRT; 397 AA.
 AC Q64194;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Lysosomal acid lipase/cholesteryl ester hydrolase precursor
 DE (EC 3.1.1.13) (LAL) (Acid cholesteryl ester hydrolase) (Sterol
 DE esterase) (Lipase A) (Cholesteryl esterase).
 GN LIPA OR LAL.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;

RX MEDLINE=96129534; PubMed=8576647;
 RA Nakagawa H., Matsubara S., Kuriyama M., Yoshidome H., Fujiyama J.,
 RA Yoshida H., Osame M.;
 RT "Cloning of rat lysosomal acid lipase cDNA and identification of the
 mutation in the rat model of Wolman's disease.";
 RL J. Lipid Res. 36:2212-2218(1995).
 CC -!- FUNCTION: CRUCIAL FOR THE INTRACELLULAR HYDROLYSIS OF CHOLESTERYL
 CC ESTERS AND TRIGLYCERIDES THAT HAVE BEEN INTERNALIZED VIA RECEPTOR-
 CC MEDIATED ENDOCYTOSIS OF LIPOPROTEIN PARTICLES. IMPORTANT IN
 CC SUPPRESSION OF HYDROXYMETHYLGLUTARYL-COA REDUCTASE AND ACTIVATION
 CC OF ENDOGENOUS CELLULAR CHOLESTERYL ESTER FORMATION (BY
 CC similarity).
 CC -!- CATALYTIC ACTIVITY: A sterol ester + H(2)O = a sterol + a fatty
 CC acid.
 CC -!- SUBCELLULAR LOCATION: Lysosomal.
 CC -!- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; S81497; AAB36043.2; -;
 DR InterPro; IPR000073; A/b hydrolase.
 DR InterPro; IPR000734; Lipase.
 DR InterPro; IPR000379; Ser esters site.
 DR Pfam; PF04083; abhydro_lipase; 1.
 DR Pfam; PF00561; abhydro_lipase; 1.
 DR PROSITE; PS00120; LIPASE SER; PALSC NEG.
 KW Hydrolase; Lipid degradation; Glycoprotein; Signal; Lysosome.
 FT SIGNAL 1 19
 FT CHAIN 20 397
 FT
 FT ACT SITE 172 372
 FT ACT SITE 372 372
 FT CARBOHYD 34 34
 FT CARBOHYD 99 99
 FT CARBOHYD 159 159
 FT CARBOHYD 271 271
 FT CARBOHYD 319 319
 FT SEQUENCE 397 AA; 45186 MW; 97A38595A0523947 CRC64;
 Query Match 54.3%; Score 1220.5; DB 1; Length 397;
 Best Local Similarity 58.2%; Pred. No. 2.7e-97;
 Matches 220; Conservative 67; Mismatches 84; Indels 7; Gaps 3;
 QY 37 PT---KAVDPEAFMNSIIHQHGPCEYEYVATEDGHILSVNRIPLGVPKTKTSRVP 93
 DB 20 PGTTSAYDPEANMNVTEIIMHWGYP--EHSVQTGDGVLGVHPIPHGRKNQFPDKGPKP 77
 QY 94 VLQV--HGLVGASWISNLPNSLGFILADAGFDVWNGSRGNWRSKHTLSIDODEF 151
 DB 78 VYLQWHEFLADSSNWNINNSLGFILADAGFDVWNGSRGNWRSKHTLSVQDEY 137
 QY 152 WAFSDPEAKYDLPASINYLKTKQEQQLXNVHSGQCTIGFIAPSQMPKELAKVKNFFA 197
 DB 138 WAFSDPEAKYDLPASINYLKTKQEQQLXNVHSGQCTIGFIAPSQMPKELAKVKNFFA 197
 QY 212 LAPATVHKASPGKELLPLDMIKGLFGKKELYQTRFLQRLVYLCOGVIDLQCSN 271
 DB 198 LAPVLISLAFSGPMYKGLRPLDLLLEDFGQKQFLPQSAWYKWLSTHICTHIVNKLCA 257
 QY 272 IMLLGGFNTNNMNSRVSVAHNTLACTSVQNLHWSQVNSGELRAFDSGSKTKLEK 331
 DB 258 IFPLCGFNEKNLNSRVDVYTHCPAGTSVQNMVHTVQVVKYHKLQAFDWSGSDKNVYH 317
 QY 332 CNOPTFVYRVDMVTPTAMVTGGQDWLSNEDVQKLLSEVTNLYHKNIPEWHAHVDPIW 391
 DB 318 YNQSPYPLYSIKMQLPTALMSGGKQWLADTSINILLTEIPTLVYHKNIPEWHDLPFIW 377

QY 392 GLDAPHRMNYNEIHLMQQ 409
 DB 378 GLDAPWRLYNEVSLMKK 395

RESULT 3

LICH MOUSE
 ID LICH_MOUSE STANDARD; PRT; 397 AA.
 AC Q9Z0M5;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Lysosomal acid lipase/cholesterol ester hydrolase precursor
 DE (SC 3.1.1.13) (LAL) (Acid cholesterol ester hydrolase) (Sterol
 DE esterase) (Lipase A) (Cholesteryl esterase).
 GN LIPA OR LIPI.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6 X CBA; TISSUE=Liver;
 RX MEDLINE=96363957; PubMed=8725147;
 RA Du H., Witte D.P., Grabowski G.A.;
 RT "Tissue and cellular specific expression of murine lysosomal acid
 RT lipase mRNA and protein";
 RL J. Lipid Res. 37:937-949(1996).
 CC -!- FUNCTION: CRUCIAL FOR THE INTRACELLULAR HYDROLYSIS OF CHOLESTERYL
 CC ESTERS AND TRIGLYCERIDES THAT HAVE BEEN INTERNALIZED VIA RECEPTOR-
 CC MEDIATED ENDOCYTOSIS OF LIPOPROTEIN PARTICLES. IMPORTANT IN
 CC SUPPRESSION OF HYDROXYMETHYLGLUTARYL-COA REDUCTASE AND ACTIVATION
 CC OF ENDOGENOUS CELLULAR CHOLESTERYL ESTER FORMATION (BY
 CC similarity).
 CC -!- CATALYTIC ACTIVITY: A sterol ester + H(2)O = a sterol + a fatty
 CC acid.
 CC -!- SUBCELLULAR LOCATION: Lysosomal.
 CC -!- TISSUE SPECIFICITY: Expressed at low levels in most tissues. High
 CC level expression is found in hepatocytes and splenic and thymic
 CC cells. Very high level expression is observed in cells of the
 CC small intestinal villi, the zona fasciculata and reticularis of
 CC the adrenal cortex, pancreatic acini, and renal tubular
 CC epithelium.
 CC -!- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; Z31689; CAA83494.1; -;
 DR MGD; MGI:96789; Lipi.
 DR InterPro; IPR000073; A/b hydrolase.
 DR InterPro; IPR000734; Lipase.
 DR InterPro; IPR000379; Ser esters site.
 DR Pfam; PF04083; abhydro_lipase; 1.
 DR Pfam; PF00561; abhydro_lipase; 1.
 DR PROSITE; PS00120; LIPASE SER; 1.
 DR Hydrolase; Lipid degradation; Glycoprotein; Signal; Lysosome.
 KW SIGNAL 1 19
 FT CHAIN 20 397
 FT
 FT ACT SITE 172 372
 FT ACT SITE 372 372
 FT CARBOHYD 34 34
 FT CARBOHYD 99 99
 FT CARBOHYD 159 159
 FT CARBOHYD 271 271
 FT SEQUENCE 397 AA; 45186 MW; 97A38595A0523947 CRC64;

```

DR PIR; A23045; LINTT.
DR InterPro; IPR000073; A/b_hydrolase.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser_estra_site.
DR Pfam; PF04083; abhydro_lipase; 1.
DR Pfam; PF00561; abhydrolase; 1.
DR PROSITE; PS00120; LIPASE_SER; 1.
KW Hydrolase; lipid degradation; Glycoprotein; Signal.
FT SIGNAL          1      18
FT CHAIN           19     395
FT ACT_SITE       171    171   TRIACYLGLYCEROL_LIPASE, LINGUAL.
FT ACT_SITE       342    342   CHARGE_RELAY_SYSTEM (BY SIMILARITY).
FT ACT_SITE       371    371   CHARGE_RELAY_SYSTEM (BY SIMILARITY).
FT DISULFID       245    254   BY_SIMILARITY.
FT CARBOHYD       33     33   N-LINKED_GLCNAC. . . (POTENTIAL).
FT CARBOHYD       68     68   N-LINKED_GLCNAC. . . (POTENTIAL).
FT CARBOHYD       98     98   N-LINKED_GLCNAC. . . (POTENTIAL).
FT CARBOHYD       270    270   N-LINKED_GLCNAC. . . (POTENTIAL).
SQ SEQUENCE        395 AA; E60185A923522EA CRC64;

Query Match               51.8%; Score 1164.5; DB 1; Length 395;
Best Local Similarity     54.9%; Pred. No. 1.7e-92;
Matches 217; Conservative 69; Mismatches 108; Indels 1; Gaps 1;

QY 17 MWLLILVAVNFQ-RNVNSVMHPTKAVDPFAFNISIEIIHQGYPCEEYEVEDGYILSV 75
DB 1 MWLLIITSVISITFGGAHLGFKLGCPNPENNWNISQMITYNGYPCEVEVVTDGYILGV 60

QY 76 NRIPRGLVQPCKTGSRPVVLLOHGLVGGASWISNLNNNISLGFILADAGDFVAMGRGN 135
DB 61 YRIPIHGKNSENGTKGPVPVYLQGHGTASNIAFLPNNSLAFLMADAGDYDVLTGNRGN 120

QY 136 AWSRKHKHTLSIDODEFWARSYDEMARPDLPAVINFIQTQEKIYYVGYSQGTTMGFIA 195
DB 121 TWSRGNYVSPDSVFETWASFDEMAKYDLPATINIVQKTGOEKLIHYVGHSGQTTIGFIA 180

QY 196 PSTMPLEAQIKMYPALAPIATVIYKAHSPTGTFELLPPDMTKELFGKEFYIQTRFLROL 255
DB 181 PSTNPTLKAKTIKFYAALPVAIVVKYTQSPLKKISPFIPTFLFKLMFGKMFLPHYTFDDFL 240

QY 256 VIVLCGVILDIOICSNIMLLLGFTNTNMMSRASVAAHTLAGTSVQNILLHNSQAENVSG 315
DB 241 GTVCSEVIDLDICSNTLIFCGFDKGNLNVSFPVILGENPAGTSVDQFLHWACLVRSG 300

QY 316 ELRAFDRGSSTKNLEKCNOPTPYRVYRDVTMTPTAMWTGGODMTLSNPEDVNKLMLSEVNTL 375
DB 301 KFOAFNWGPSQNMHLHYNQKPTEPDVSANTVPVAVNGNDILLADPODVAMLPLKLSNL 360

QY 376 IYKHNIPEWAHVDFIWCLDAPEHMYNEIHLMQQE 410
DB 361 LPFHKEILAYNHLDFTWMDAPAEVYNEMISMMAED 395

RESULT 5
LIPG CANFA STANDARD; PROT; 398 AA.
ID LIPG CANFA
AC P80035; O02857;
DT 01-NOV-1991 (Rel. 20, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Triacylglycerol lipase, gastric precursor (EC 3.1.1.3) (Gastric
DE lipase) (GL).
DN LIFF.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Vertebrate; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Pissepedia; Canidae; Canis.
OX NCBI_TaxId=9615;
RN [1]_
RN RZ SEQUENCE FROM N.A.
RP TISSUE=Stomach;
RX MEDLINE=99450174; PubMed=10520456;
RA Vaganay S., Joffé G., Bertaux O., Toselli E., Devignes M.D.,
RA Benicourt C.;
```

"The complete cDNA sequence encoding dog gastric lipase.";
 [2]

```

RT DNA Seq. 8:57-262(1998).
RN SEQUENCE OF 20-59.
RX MEDLINE=92037652; PubMed=1935982;
RA Carriere F., Moreau H., Raphael V., Laugier R., Benicourt C.,
RA Junien J.-L., Verger R.;
RT "Purification and biochemical characterization of dog gastric
RT lipase.";
RL Eur. J. Biochem. 202:75-83(1991).
CC -1- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O = diacylglycerol + a
CC fatty acid anion.
CC -1- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR PDB; 1X8Q; 20-MAR-02.
DR InterPro; IPR000073; A/b hydrolase.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser_estr_1.
DR Pfam; PF04083; abhydro_lipase; 1.
DR Pfam; PF00561; abhydro_lipase; 1.
DR PROSITE; PS00120; LIPASE SER; 1.
KW Hydrolase; Lipid degradation; Glycoprotein; Signal; 3D-structure.
FT SIGNAL 1 19
FT CHAIN 20 398 TRIACYLGLYCEROL LIPASE, GASTRIC.
FT ACT_SITE 172 372 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 372 372 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CARBOHYD 34 34 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 39 39 I -> T (IN REF. 2).
SQ SEQUENCE 398 AA; 45130 MW; E04D62F7518386C CRC64;
Query Match 51.4%; Score 1154; DB 1; Length 398;
Best Local Similarity 54.3%; Pred. No. 1.4e-91;
Matches 216; Conservative 71; Mismatches 99; Indels 12; Gaps 5;
QY 17 MWLLILVAFMFORNVSVH-----MPTKAVDPFAFNISLIHQHGYPCBEYEVATEDG 70
Db 1 MWLLTAAASVIS-TIGTTHGLFGKLUHPT---NPEVTNNISQMITTWGYPABEYEVATEDG 56
QY 71 YILSVNRIIPRGVLQPKTGSRPVLLQHLGVGASNWTISNLPNNSLGFLADAGFDVWVG 130
Db 57 YILGIDRIPIYGRKNSNIGRRPVAFLQHLGLASATNWTISNLPNNSLAPILADAGVDVWLG 116
QY 131 NSRGNAWKKHTLIDQDEFWAFSYDEMAREFDLPVINFILQKTQBEKIYVGSQGT 190
Db 117 NSRGNTWARNDIYSPDSVEFWAFSDEMAYDLPATIDFLIKTQDQKLHYVGSQGT 176
QY 191 MGFIAFTMPELAQKIMVAFALAPATVTKHAKSPGTEKLLLPDMNIGLFGKBEFLYQTR 250
Db 177 IGFIAFTNPFLAKIKFYALAPATVTKYETLNLKMLVPSFLKIFGNKTF-YPHH 235
QY 251 FLRQ-LVLYCQGVLDICSNIMLLGFTNTNNMNSRASVYAHTLAGTSVQNILHWS 309
Db 236 FFDQFLATEVCSSRETVLLCSNALPTICGFTDMLNMSRLDYLSHNPAGTSVQNVLHWS 295
QY 310 QAVNSELRAPDWSGETNLEKCNQPTVRYEVDRDWTPTAMWTGGQDLNLPDSDVKMLL 369
Db 296 QAVKSGKQAFDWSGTVQNMHYSMPYPYNTLDMHVFIAVWNGNDLLADPHDVL 355
QY 370 SEVTNLIYHKNPPEWAHVDYFWGLDAPHRMNTIETHLM 407
Db 356 SKLPNLIVYHKRIPFPYHLDYFWAMPDAPQAVNYEIVSM 393

```

RESULT 6

```

LIPG BOVIN STANDARD; PRT; 397 AA.
AC Q29458;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Triacylglycerol lipase, pregastric precursor (EC 3.1.1.3) (Pregastric
DE lipase) (GL) (Pregastric esterase) (PGE).
GN LIPF.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE= Tongue serous gland;
RX MEDLINE=95011625; PubMed=7926811;
RA Timmermans M.Y.J., Kupers L.P., Teuchy H.;
RL "The cDNA sequence encoding bovine pregastric esterase.";
RL Gene 147:259-262(1994).
CC -1- CATALYTIC ACTIVITY: Triacylglycerol + H(2)O = diacylglycerol + a
CC fatty acid anion.
CC -1- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; L26319; AAA57037.1; -
PIR; JC4017; JC4017.
InterPro; IPR000073; A/b hydrolase.
InterPro; IPR000734; Lipase.
InterPro; IPR000379; Ser_estr_1.
Pfam; PF04083; abhydro_lipase; 1.
Pfam; PF00561; abhydro_lipase; 1.
DR PROSITE; PS00120; LIPASE SER; 1.
KW Hydrolase; Lipid degradation; Glycoprotein; Signal.
FT SIGNAL 1 19 BY SIMILARITY.
FT CHAIN 20 397 TRIACYLGLYCEROL LIPASE, PREGASTRIC.
FT ACT_SITE 171 371 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 371 371 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CARBOHYD 33 33 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 270 270 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 326 326 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 397 AA; 45231 MW; F69977DED58EE36 CRC64;
Query Match 50.7%; Score 1139; DB 1; Length 397;
Best Local Similarity 53.7%; Pred. No. 2.7e-90;
Matches 212; Conservative 68; Mismatches 113; Indels 2; Gaps 2;
QY 18 WLLILVAFMFORNVSVH-----MPTKAVDPFAFNISLIHQHGYPCBEYEVATEDGYLSVNR 77
Db 3 WLLVTVCFIHMGSNAFCFLGKIAKAPASMNYSQMSIYWGYPSEMHKVTADGYLQVYR 62
QY 78 IPRGLVQPKTGSRPVLLQHLGVGASNWTISNLPNNSLGFLADAGFDVWNGSRGNW 137
Db 63 IPRGNKNNHNLQORPVVFLQHLGLASATNWTISNLPNNSLGFLADAGFDVWNLGSRGNTW 122
QY 138 SRKHKTLSDQDEFWAFSYDEMAREFDLPVINFILQKTQBEKIYVGSQGTMTGFIAPS 197
Db 123 AQEHLTYSDSEFWAFSDEMAREFDLPSTIDFLIKTQDQKLHYVGSQGTITGFIAPS 182
QY 198 TWPPELAQKIMVAFALAPATVTKHAKSPGTEKLLLPDMNIGLFGKBEFLYQTRFLRQ-IV 256
Db 183 TSPTLAEKIKVFYALAPATVTKYSLFNKLALIPHFLPKIIFGDKMF-YPHTFLEQLG 241

```

GO; GO:0004806; P:triacylglycerol lipase activity; TAS.	
GO; GO:0006441; P:triacylglycerol metabolism; NAS.	
InterPro; IPR000073; A/b hydrolase.	
InterPro; IPR000734; Lipase.	
InterPro; IPR000379; Ser esters site.	
Pfam; PF04083; abhydro lipase; 1.	
Pfam; PF00561; abhydro lipase; 1.	
PROSITE; PS00120; LIPASE_SER; 1.	
KW Hydrolase; lipid degradation; Glycoprotein; Signal; Polymorphism;	
3D-structure.	
SIGNAL	1 19
CHAIN	20 398
FT ACT_SITE	172 172 TRIACYLGLYCEROL LIPASE, GASTRIC.
FT ACT_SITE	343 343 CHARGE RELAY SYSTEM.
FT ACT_SITE	372 372 CHARGE RELAY SYSTEM.
FT DISULFID	246 255
FT CARBOHYD	34 34 N-LINKED (GLCNAC. .).
FT CARBOHYD	99 99 N-LINKED (GLCNAC. .).
FT CARBOHYD	271 271 N-LINKED (GLCNAC. .).
FT CARBOHYD	327 327 N-LINKED (GLCNAC. .).
FT VARIANT	161 161 T -> A (IN dbsNP:814628).
SEQUENCE	398 AA; 45237 MW; CDSEEL62LCIL4FOF CRC64;
Query Match	49.8%; Score 1119; DB 1; Length 398;
Best Local Similarity	52.2%; Pred. No. 1.4e-88;
Matches 210; Conservative	73; Mismatches 107; Indels 12; Gaps 2;
QY	17 MWLLIIVAYMFQRNVNSVMPKAV-----DPEAFNWNISEIIHQGCPCEEVEVATED 69
DB	1 MWLLLTWASLI-----SVLGTTGHLFGKLHPGSPEVTWNISQMITYGYPMEEYEVTVD 55
QY	70 GYLIVSNRIEPLGVLPVKTKGRSPVVLLHQLVGCGASNIMSLPNNSLGFILLADAGFDVWM 129
DB	56 GYLIEVNRIPYGKNSGNTGQRPVVFLQGLLASATNMISLNPNNSLAIFILLADAGYDWL 115
QY	130 GNSRGNVAWRKHKTLSIDQDEFWAPSYDEMARFDLPVINFILOKTGEKIYYVYGSGT 189
DB	116 GNSRGNTWARNRLYSPDSVEFWAFSDFEMAKYDLPATIDFIVCKTGOKQLHYVGHSGT 175
QY	190 TMGTFASTWPELLAAOKIKMYFALLAPTATVKHAKSPTGTFLLPDMWKGLFKGKEFLYT 249
DB	176 TIGTIASTWPSLAIRKIFYALAPVATVYKYSLINKLRFPVQSLEFXIFGDKIFYPHN 235
QY	250 RFLRLQVLYICGGVILDQCISNIMLLGFGNTNNMNGRASVYAAHTLAGTSVQNILHWS 309
DB	236 FFDQFLATEVCRESRMNLCSNALFITICGFSKNFNSTRSLDYLSHNPAGSTSQVMHFMT 295
QY	310 QAVNSGELRAFADGSETKNLEKNCNQTPVRYVRDMVTPTAMWTGGQDWLNPEPDVKML 369
DB	296 QAVXSGKFQAYDWMSQPONRWHDQSQPPYNTANVFIAVNWGKKOLLADPDQDVGLLL 355
QY	370 SEVNTNLIYHKNIPEWAHVDFINGLDAPHMYNEIHLMQOE 411
DB	356 PKGLPNLIYHKEIPYNELDFIWMDAQEVYNDIVSMISEDK 397
RESULT 8	
LIP3 DROME	
ID LIP3 DROME STANDARD; PRZ; 394 AA.	
AC O46108;	
DT 16-OCT-2001 (Rel. 40, Created)	
DT 16-OCT-2001 (Rel. 40, Last sequence update)	
DT 16-OCT-2001 (Rel. 40, Last annotation update)	
DE Lipase 3 precursor (EC 3.1.1.-) (Dmlip3).	
GN LIP3 OR CG9823.	
OS Drosophila melanogaster (Fruit fly).	
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;	
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;	
OC Ephydroidea; Drosophilidae; drosophila.	
NCBI_Taxid=7227;	
[1]	
RN	SEQUENCE FROM N.A.
RP	

[illegible]

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.G., Ferraz C., Ferrier S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glaser K.,
RA Glodak A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jatali M., Kaulush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler P., Shen H.,
RA Shue B.C., Sriden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spalding A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*;"
RL Science 287:2185-2195 (2000).
[3]
RN SEQUENCE FROM N.A.
RC STRAIN=Berkeley; TISSUE=Embryo;
RX MEDLINE=22426066; PubMed=12537569;
RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA George R.A., Guarin H., Krommiller B., Pacle J.M., Park S., Wan K.H.,
RA Rubin G.M., Celniker S.E.;
RT "A *Drosophila* full-length cDNA resource;"
RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8 (2002).
CC -!- FUNCTION: COULD BE A DIGESTIVE ENZYME.
CC -!- SUBCELLULAR LOCATION: Secreted (Probable).
CC -!- TISSUE SPECIFICITY: In 14 hour embryos expression is seen in the
CC foregut/midgut boundary.
CC -!- DEVELOPMENTAL STAGE: Expressed from 14 hour embryos through to
CC adulthood. There is a weak maternal contribution to early embryos.
CC -!- SIMILARITY: PARTIAL WITH OTHER LIPASES (PANCREATIC, GASTRIC,
CC HEPATIC, LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.). ALSO SIMILAR TO
CC LIPIDOPROTEIN EGG-SPECIFIC AND YOLK PROTEINS.
CC -!- CAUTION: It is uncertain whether Met-1 or Met-7 is the initiator.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC of the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y14366; CAA74736.1; ALT_INIT.
CC EMBL; A003629; AAF52994.1; -
CC EMBL; AY075506; AAL68315.1; -
CC Flybase; FBgn0023496; Lipi.
CC InterPro; IPR000073; A/b hydrolase.
CC InterPro; IPR000734; Lipase.
CC InterPro; IPR000379; Ser esters site.
CC Pfam; PF04083; abhydro lipase; 1.
CC Pfam; PF00561; abhydro lipase; 1.
CC PROSITE; PS00120; LIPASE_SER; FALSE_NEG.
CC HydroLase; Lipid degradation; Signal; Glycoprotein.
CC SIGNAL 1 24 POTENTIAL.
CC CHAIN 25 439 LIPASE 1.
CC DOMAIN 30 44 POLY-GLU.

FT ACT_SITE 197 197 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 393 393 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CARBOHYD 124 124 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 346 346 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 379 379 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 426 426 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 10 10 I -> I (IN REF. 1).
FT CONFLICT 213 213 Y -> F (IN REF. 1).
FT CONFLICT 412 412 Q -> E (IN REF. 1).
SQ SEQUENCE 439 AA; 50660 MW; 9E32E20BEAE93E3F CRC64;
Query Match 24.3%; Score 546; DB 1; Length 439;
Best Local Similarity 31.8%; Pred. No. 2.8e-39;
Matches 123; Conservative 93; Mismatches 157; Indels 24; Gaps 7;
Qy 39 KAVDPEAFNMSIEIIHQHQPCEEYEVATEDGYILSVNRIPLGLVQPKTKGRPVVLLQH 98
Db 58 KNIKQDSTLSVDKLIAGYGVSEVHVHTTDDGYILTMHRI-----RKQGA-PPFLLOH 109
Qy 99 GLVGGASNWSINLPNNSLGFIILADAGFDVWNGSRGNAMSRKHKTLISIDQDEFWAFSYDE 158
Db 110 GLVDSSAGFVGMGNFVSLAYLLADHNDVWLGNGRYSRNHTLLDPDSKFWDSWIE 169
Qy 159 MARFDPAVINFLQTKGQEKIYYGVSGQTTMGFTAFSTMPPELAQIKIKYFALAPIATV 218
Db 170 IGMVLDLPAMIDHVLKVTGPKLHVAGHSQCTSFVWCSMRPAYNDKVYVSMQALAPAVYA 229
Qy 219 KHAKS-PGTGKLLPDMWIKGLFGKKEFLYQTRFLRQLVYLGGVILDDQISNIMLLG 277
Db 230 KETEDHPYTRAIISLYNSLVGSSIREMFNGEFRFLCRMT-----EETELCLIEAVFGIV 283
Qy 278 GFNTNNMNSRASVYAAHTLAGTSVQNIHLWSQAVNSGELRAPDWGSETKNLEKCNQPTP 337
Db 284 GRWNEFNRCWFFVILGHYPAGVAAKQVHFQIKSGRFAPVSY-SSKNQMQLYRDHLP 342
Qy 338 VRYRVRDMVTPTAMWTGGQDLNSPDDVQMLSEVNTLYLHKNIIP-EWAHVDVFIWGLDA 395
Db 343 PRYNLSLVTVPTFTVYVYSTNDLLCHPKDVFESMCDLGNVTGKYLVPQKEFNHMDFLMAIDV 402
Qy 396 PHRYNNEIIHLM-----QOEETNLSQ 416
Db 403 RMLYRRMLQVLGKVPESPEANRSR 429
RESULT 10
TGL1 YEAST STANDARD; PRT; 548 AA.
AC P34163;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Triglyceride lipase-cholesterol esterase (EC 3.1.1.-).
GN TGL1 OR YKL140W OR YKL5.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92245761; PubMed=1574929;
RA Abraham P.R., Mulder A., Van't Riet J., Planta R.J., Raue H.A.;
RT "Molecular cloning and physical analysis of an 8.2 kb segment of
RT chromosome XI of *Saccharomyces cerevisiae* reveals five tightly linked
RT genes;"
RL Yeast 8:227-238 (1992).
RN [2]
RP SEQUENCE FROM N.A.
RA Rad M.R., Xu G., Kirschner L., Fritz C., Keuchel H., Hollenberg C.P.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the AB hydrolase superfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

RESULT 11	YD94	METJA	STANDARD;	PET;	987 AA.
ID	YD94	METJA			
AC	Q58789;				
AD	15-JUL-1998	(Rel. 36, Created)			
DT	15-JUL-1998	(Rel. 36, Last sequence update)			
DT	28-FEB-2003	(Rel. 41, Last annotation update)			
DE	Hypothetical protein MJ1394.				
GN	MJ1394.				
OS	Methanococcus jannaschii.				
OC	Archaea; Euryarchaeota; Methanococci; Methanococcales;				
OC	Methanocaldococcaceae; Methanocaldococcus.				
OX	NCBI_TaxID=2190;				
[1]	_RN				
RC	SEQUENCE FROM N.A.				
RP	STRAIN=JAL-1 / DSM 2661 / ATCC 43067;				
RY	MEDLINE=96337999; PubMed=8688087;				
RY	Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,				
RA	Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,				
RA	Kerlavage A.R., Dougherty B., Tomb J.-F., Adams M.D., Reich C.I.,				
RA	Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,				

RESULT 12			
ORC2_CABEL		ORC2_CABEL	
ID	ORC2_CABEL	STANDARD;	PRT; 430 AA.
AC	Q21037;		
DT	01-NOV-1997	(Rel. 35, Created)	
DT	15-JUL-1998	(Rel. 36, Last sequence update)	
DT	28-FEB-2003	(Rel. 41, Last annotation update)	
DE	DE	Origin recognition complex subunit 2 (CeOrc2)	
GN	ORC-2 OR F59E10.1.		
OS	Ceenorhabditis elegans.		

RESULT 13	
PAPA_CAVPO	
ID PAPA_CAVPO	STANDARD;
AC P70853;	PRT; 436 AA.
DT 01-NOV-1997	(Rel. 35, Created)

```

120 LKLLYGVKVPKAWNSPLTKGKYPLIFSHGL--GAFRSI---YSAIGIELASHGFI 173
      : : : : : : : : : : : : : : : : : : : : : : : : : : :
127 V-----WMSGNSRGAWSRKHKTLSDODE-----FW 152

174 VAAVEHRDESAATYYFODAPAAESGNRWIYYKVNLETEBTKRQLFORCEGSOALS 213
      : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

QY 153 AFSYDEMARDPDAVINFILOK-----TGOEKIYYVYSGQTMGFIAFTMPPELAQIKM 208
Db 234 LUSIDGEPVKNVLDNFIQOGLSLDRSKVAIGHSGF---GATVITQTLSE-DQRFRC 289
QY 209 YFAL-----APIATVHXAKSPGKFKLLPDMIMKGLFGKK-BFLYQTRFLRLQV----- 256
Db 290 GIALDPMPFVGVGDEYHVKIPQPLFINSYFOSANDTKIEKFYQPKERKXIAVKGSVH 349
QY 257 -----IYLCQGVILDOICSNMILLGGFNTNNMWSRASVYAAHTLAGTSVQNTLHWSQ 310
Db 350 HNFVDFTFATGKII-----GQMLSKGKIDSEVAMDINKASLAFLOKYLGLD-KNFDQ 402
QY 308 WSOAVNSGELAFDWSGSETKNLEKCNQPTPVRYRVRDVTPTAMTGGQDWLSNPPE 363
Db 403 WNSLME-----GDDENL-----IPEFTIPTMQSSGTGTQRNPD 436

RESULT 14
Y193 HAEIN
ID Y193 HAEIN STANDARD; PRT; 287 AA.
AC Q57427; O05013;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative esterase/libase HI0193 (EC 3.1.1.-).
GN HI0193.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-P., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512 (1995).
CC -!- SIMILARITY: BELONGS TO A FAMILY OF ESTERASES THAT GROUPS TOGETHER
CC PSEUDOMONA TROPINESTERASE, DMPD, TODF AND XYLE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U32704; AAC21862.1; -
CC DR PIR; E64053; E64053.
CC TIGR; HI0193; -
CC InterPro; IPR000073; A/b hydrolase.
CC InterPro; IPR003089; AB hydrolase.
CC InterPro; IPR000379; Ser_estrases.
CC Pfam; PF00561; abhydrolase; 1.
CC PRINTS; PRO0111; ABHYDROLASE.
CC KW Hypothetical protein; Hydrolase; Serine esterase; Complete proteome.
FT ACT SITE 119 BY SIMILARITY.
FT ACT SITE 266 266 BY SIMILARITY.
SQ SEQUENCE 287 AA; 32987 MW; F28548619C060619 CRC64;

Query Match 4.3%; Score 96.5; DB 1; Length 287;
Best Local Similarity 19.1%; Pred. No. 0.67;

```

```

Matches 62; Conservative 39; Mismatches 119; Indels 105; Gaps 12;
QY 84 QPKKTSRPPVLLQHLGVGASNWISNPNNSLGLTADAGDFVMNGSRGNAMSKKHT 143
Db 39 QVQKQTINTPVLITPHGLGDMON-----LGVI-----ARAFSEHYST 75
QY 144 LSIQDQDEFWAFSDEMARFDLPAV-INFILQKTQGEKIYYVYSGQTMGFIAFTMPPEL 202
Db 76 LRIDLNRHGHGSFSEKKNYQLMAEDVIAVRHNLNLSKVILGHSMGCKTAMKITALCPBL 135
QY 203 AQIKMYFALAPIATYVHXAKSPGKFKLLPDM-----MIKGLFGKKEFLYQTR 250
Db 136 VEK-----LVIDNSMPYEGFGHKDVFNGFLFAVKNAPEN- 171
QY 251 FLRLVLYLGGVILDOICSNMILLGGFNTNNMWSRASVYAAHTLAGTSVQNTLHWSQ 310
Db 172 --RQAKPILKQEIENDE--DVQFMKLSFDVNSADCFRNL-----TALFNNTANINMDE- 222
QY 311 AVNSGELAFDWSGSETKNLEKCNQPTPVRYRVRDVTPTAMTGGQDWLSNPPE 370
Db 223 -----KVRVF-----TPTLFIKGSSSVYIKIENSEKILE 251
QY 371 EVNLIYH--KNIPFWAHV---DPI 390
Db 252 OFFNATAFTINGSGHWVHAEPDFV 276

RESULT 15
L1N1 HUMAN
ID L1N1 HUMAN STANDARD; PRT; 1259 AA.
AC P08547;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 01-AUG-1988 (Rel. 08, Last annotation update)
DE LINE-1 reverse transcriptase homolog.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE.
RX MEDLINE=66230917; PubMed=2423883;
RA Hattori M., Kuhara S., Takenaka O., Sakaki Y.;
RT "A family of repetitive DNA sequences in primates may be derived
RT from a sequence encoding a reverse transcriptase-related protein.";
RL Nature 321:625-628 (1986).
CC -!- MISCELLANEOUS: THIS SEQUENCE WAS CONSTRUCTED FROM AN ALIGNMENT OF
CC PUBLISHED AND UNPUBLISHED SEQUENCES, DETERMINED IN VARIOUS
CC LABORATORIES, BELONGING TO THE LINE-1 FAMILY.
CC PIR; A25313; GHUUL.
CC DR InterPro; IPR005135; Exo_endo_phos.
CC DR InterPro; IPR000477; RVTse.
CC DR Pfam; PF03372; Exo_endo_phos; 1.
CC DR Pfam; PF00078; rvt; 1.
CC KW RNA-directed DNA polymerase.
SQ SEQUENCE 1259 AA; 147112 MW; F3BC18A3803919F5 CRC64;

Query Match 4.3%; Score 95.5; DB 1; Length 1259;
Best Local Similarity 19.7%; Pred. No. 5.9;
Matches 90; Conservative 67; Mismatches 170; Indels 129; Gaps 23;
QY 28 QRNVNSVHMTKAVDPEAFNFI---SEIIHQGYPCPEYE----- 64
Db 442 QEEVESLNRPITSSEIEALINSPLNPKSPQEPGFTAEFYQRYKEELVPFLKLQSIKKE 501
QY 65 -VATEDGYILSVNRIPR-GLVQPKTKGSRPVL-----LQHLVGGASNWSNL-PNN 114
Db 502 GILPNSFYEAISILIPKGRDRTTKENFRPISLNNIDAKILNLANQIQOHKIKLIHD 561
QY 115 SLGFIADAGDFVMNGSR-----GNAMSKKHTLSIDQDEFWAFSYDEM----- 159
Db 562 QVGFIPAMQG---WFNIRKSINIIOHNRKOTNHNMIISDAEK-----AFDKIQOFFMLK 614

```

```

Qy 160 -----ARFDPVINEILOKTOEKIYYVYGSQGTMGFIASFMPELA 203
Db 615 PLNKLGDGYLYKIIRAIYDKPTA-NIIL--NGQ-KLEAPPLKTGTROGCPSPLEPNIV 670
Qy 204 QKIMWYFALAPIATVVKHAKSPGTFKLLPDMMIKGLFGKKEFLYQTRFLRQLVIVYLCGOV 263
Db 671 LEV-----LA--RAIROKE-----IKGIQKGKEEVKLSLFADDMMIVYLENFI 711
Qy 264 ILDOICSNIMLLGGFN-----TNNMMSPASVYAAHTLAGTSVQNIHWSQVNSGELR 318
Db 712 VSAQ---NLKLSNFSKVSQYKINVKSOAFLYTNNRQTESQIMSELPFTIASKRIKYL 768
Qy 319 AFDWGSSTKLEKCNQPTPVRYVRDWT-----VPTAMWTGGQDMLSNPEDVKMLISEVT 373
Db 769 GIQTRDVKDLFKENY-KPLNEIKEDTNKKNIPCS-----WVGRINIVKMAI--LP 818
Qy 374 NLIYHKN-IP-----EWAHVDPIWGLDAPH 397
Db 819 KVIYRFNAIPKLPMTFFTELEKTTLKFIVNQKRAH 854

```

Search completed: February 19, 2004, 14:21:11
Job time : 14.8856 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2004, 14:17:14 : Search time 30.1771 Seconds
(without alignments)
3617.186 Million cell updates/sec

Title: US-10-042-431-47
Perfect score: 2247

Sequence: 1 MLETLRQWIVSRHMEKWL.....IHLMQEETNLQGRCEAVL 423

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1991	88.6	374	4	Q96LG2
2	1975.5	87.9	422	11	Q8K2A6
3	1971.5	87.7	422	11	Q8BJ28
4	1222.5	54.4	397	11	Q8C2G7
5	1134.5	50.5	395	11	Q9D798
6	1134.5	50.5	395	11	Q9D606
7	1134.5	50.5	395	11	Q9D6P3
8	1134.5	50.5	395	11	Q9D6X0
9	1133.5	50.4	395	11	Q9CPB7
10	1133.5	50.4	395	11	Q9CPB8
11	1131.5	50.4	395	11	Q9D6L9
12	1129.5	50.3	395	11	Q9D6F5
13	1128.5	50.2	395	11	Q9D767
14	1127.5	50.2	395	11	Q9D7C5
15	1126.5	50.1	395	11	Q9D6L1
16	1126.5	50.1	395	11	Q9D766

17	1125.5	50.1	395	11	Q9D6Q3
18	1124.5	50.0	395	11	Q9D760
19	1124.5	50.0	395	11	Q9D6S5
20	1122.5	50.0	395	11	Q9D796
21	1119.5	49.8	395	11	Q9D6N8
22	1116.5	49.7	395	11	Q9D6Q2
23	1115.5	49.6	395	11	Q9D6M9
24	1098	48.9	398	11	Q8BM14
25	829.5	36.9	275	13	Q8AYB8
26	755.5	33.6	684	5	Q95XV1
27	752.5	33.5	403	5	Q61866
28	751	33.4	411	5	Q20449
29	730	32.5	405	5	Q93789
30	716	31.9	404	5	Q16956
31	692	30.8	411	5	Q94252
32	668.5	29.8	351	5	Q95X33
33	666	29.6	422	5	Q9VKT8
34	655.5	29.2	426	5	Q17766
35	655.5	29.2	434	5	Q9VKT9
36	650.5	28.9	531	5	Q9VKS5
37	636	28.3	398	5	Q9V796
38	625.5	27.8	616	5	Q77107
39	602	26.8	429	5	Q8MWX1
40	586.5	26.1	421	5	Q9VBK6
41	574.5	25.6	559	5	Q17219
42	567.5	25.3	457	5	Q9VKT2
43	567.5	25.3	457	5	Q8T3X7
44	559	24.9	399	5	Q9VPE9
45	535	23.8	416	5	Q9VQQ5

ALIGNMENTS

RESULT 1

ID	Q96LG2	PRELIMINARY;	PRT;	374 AA.
AC	Q96LG2;			
DT	01-DEC-2001 (TrEMBLrel. 19, Created)			
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)			
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)			
DE	BA30415.1 (Novel lipase) (Fragment).			
GN	BA30415.1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Bray-Allen S.;			
RL	Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.			
DR	EMBL; AL353113; CAC78754.1; -			
DR	InterPro; IPR000073; A/b hydrolase.			
DR	InterPro; IPR006693; abhydro_lipase.			
DR	InterPro; IPR000734; Lipase.			
DR	InterPro; IPR000379; Ser esters_site.			
DR	Pfam; PF00561; abhydro_lipase; 1.			
DR	PROSITE; PS004083; abhydro_lipase; 1.			
FT	NON TER			
SQ	SEQUENCE 374 AA; 42371 MW; 5E7220A889437337 CRC64;			

Query Match 88.6%; Score 1991; DB 4; Length 374;
Best Local Similarity 100.0%; Pred. No. 8.4e-174;
Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	50	SEIIHQGVPCREYEVATEDGYILSVNIRPGIVOPKKTGSRPVLLQHLGVGGASNTS	109
DB	1	SEIIHQGVPCREYEVATEDGYILSVNIRPGIVOPKKTGSRPVLLQHLGVGGASNTS	60
QY	110	NLPNNSLGFLADAGFDVVMGNSRGNWGRKHKTLISIDQDEFWAFSYDEMARFDLPVAVIN	169
DB	61	NLPNNSLGFLADAGFDVVMGNSRGNWGRKHKTLISIDQDEFWAFSYDEMARFDLPVAVIN	120

QY 170 FILQKQZQEKIYVGVSGTGTGFIATFSTMPALAKIKMYFALAPIATVHKHAKSPGKPL 229
 DB 121 FILQKQZQEKIYVGVSGTGTGFIATFSTMPALAKIKMYFALAPIATVHKHAKSPGKPL 180
 QY 230 LLPDMIKGLFGKKEFLYQTRFLRQLVILCGQVILDOICSNIMLLGGFNTNNMMSRA 289
 DB 181 LLPDMIKGLFGKKEFLYQTRFLRQLVILCGQVILDOICSNIMLLGGFNTNNMMSRA 240
 QY 290 SVYAAHTLAGTSVONILHWSQAVNSGELRAFDWGSSEKNCQPTPVRYVRDVTPT 349
 DB 241 SVYAAHTLAGTSVONILHWSQAVNSGELRAFDWGSSEKNCQPTPVRYVRDVTPT 300
 QY 350 AMWTGGQDWSNEDVKMLSEVNTLIYHKNIPEWAHVDFIWLGDAPHRMYNEIHLMOQ 409
 DB 301 AMWTGGQDWSNEDVKMLSEVNTLIYHKNIPEWAHVDFIWLGDAPHRMYNEIHLMOQ 360
 QY 410 EETNLSQRCRAVL 423
 DB 361 EETNLSQRCRAVL 374

RESULT 2

Q8K2A6 PRELIMINARY; PRT; 422 AA.
 AC Q8K2A6; 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE RIKEN CDNA 4632427C23 gene (BA30415.1).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Mammary gland, and Vagina;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 DR EMBL; BC031933; AAH31933.1; -
 DR EMBL; AK037091; BAC29699.1; -
 DR EMBL; AK085719; BAC39517.1; -
 DR InterPro; IPR000073; A/b hydrolase.
 DR InterPro; IPR006693; abhydro_lipase.
 DR InterPro; IPR000734; Lipase.
 DR InterPro; IPR000379; Ser estsr_site.
 DR Pfam; PF00561; abhydro_lipase; 1.
 DR Pfam; PF04083; abhydro_lipase; 1.
 DR PROSITE; PS00120; Lipase_SER; 1.
 SQ SEQUENCE 422 AA; 48253 MW; C6EFBC140963E051 CRC64;

Query Match 87.9%; Score 1975.5; DB 11; Length 422;
 Best Local Similarity 87.2%; Pred. No. 2.7e-172;
 Matches 369; Conservative 23; Mismatches 30; Indels 1; Gaps 1;
 QY 1 MLETLRQWIVSHRMEMWLLILVAYMFQNVNSVHMTKAVDPEAFNISEIIHQGYPC 60
 DB 1 MSELRSRVTVSHRVEIWLILVAYLLQRNVNSGHLPTKAADPEAFNVSEIIKHGYP 60
 QY 61 EYEVATEDGYLSVNRIPRGVLPKKTGSRPVVLLQHLGVGASNWSNLNNLSGLFIL 120
 DB 61 EYEVATEDGYLSVNRIPRGVLPKKTGSRPVVLLQHLGVGASNWSNLNNLSGLFIL 120
 QY 121 ADAGFDVWNGSRGNAMSRKHTLSIDQDEFWAFSYDEMARFDPVAVINFILQKTQEKI 180

DB 121 ADAGFDVWNGSRGNAMSRKHTLSIDQDEFWAFSYDEMARFDPVAVINFILQKTQEKI 180
 QY 181 YVGVSGTGTGFIATFSTMPALAKIKMYFALAPIATVHKHAKSPGKPLLPDMIKGLP 240
 DB 181 YVGVSGTGTGFIATFSTMPALAKIKMYFALAPIATVHKHAKSPGKPLLPDMIKGLP 240
 QY 241 GKKEFLYQTRFLRQLVILCGQVILDOICSNIMLLGGFNTNNMMSRASVYAAHTLAGT 300
 DB 241 GRQEFLYQTRFRQLFYILCGQVILDOICSNIIILLGGFNTNNMMSRANVYVAHTPAGT 300
 QY 301 SVQNILHWSQAVNSGELRAFDWGSSEKNCQPTPVRYVRDVTPTAMWTGGQDWSL 360
 DB 301 SVQNILHWSQAVNSGELRAFDWGSSEKNCQPTPVRYVRDVTPTAMWTGGQDWSL 360
 QY 361 NPEVDKMLLSHVNTLIYHKNIPEWAHVDFIWLGDAPHRMYNEIHLMOQSEETNLSQRCR 420
 DB 361 NPDDVKLLSEVNTLIYHKNIPEWAHVDFIWLGDAPHRMYNEIHLMOQ-EPNLPQSTCR 419
 QY 421 AVL 423
 DB 420 VKL 422

RESULT 3
 Q8BJ28 PRELIMINARY; PRT; 422 AA.
 AC Q8BJ28; 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE BA30415.1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Skin;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 DR EMBL; AK037214; BAC39757.1; -
 SQ SEQUENCE 422 AA; 48287 MW; 5F2522D0E27050C8 CRC64;

Query Match 87.7%; Score 1971.5; DB 11; Length 422;
 Best Local Similarity 87.0%; Pred. No. 6.2e-172;
 Matches 368; Conservative 23; Mismatches 31; Indels 1; Gaps 1;
 QY 1 MLETLRQWIVSHRMEMWLLILVAYMFQNVNSVHMTKAVDPEAFNISEIIHQGYPC 60
 DB 1 MSELRSRVTVSHRVEIWLILVAYLLQRNVNSGHLPTKAADPEAFNVSEIIKHGYP 60
 QY 61 EYEVATEDGYLSVNRIPRGVLPKKTGSRPVVLLQHLGVGASNWSNLNNLSGLFIL 120
 DB 61 EYEVATEDGYLSVNRIPRGVLPKKTGSRPVVLLQHLGVGASNWSNLNNLSGLFIL 120
 QY 121 ADAGFDVWNGSRGNAMSRKHTLSIDQDEFWAFSYDEMARFDPVAVINFILQKTQEKI 180
 DB 121 ADAGFDVWNGSRGNAMSRKHTLSIDQDEFWAFSYDEMARFDPVAVINFILQKTQEKI 180
 QY 181 YVGVSGTGTGFIATFSTMPALAKIKMYFALAPIATVHKHAKSPGKPLLPDMIKGLP 240
 DB 181 YVGVSGTGTGFIATFSTMPALAKIKMYFALAPIATVHKHAKSPGKPLLPDMIKGLP 240
 QY 241 GKKEFLYQTRFLRQLVILCGQVILDOICSNIMLLGGFNTNNMMSRASVYAAHTLAGT 300
 DB 241 GRQEFLYQTRFRQLFYILCGQVILDOICSNIIILLGGFNTNNMMSRANVYVAHTPAGT 300
 QY 301 SVQNILHWSQAVNSGELRAFDWGSSEKNCQPTPVRYVRDVTPTAMWTGGQDWSL 360

Db 301 SVONILHWSQAVNSGELRAFDWGTSEKQKCNQPTPIRYKYRDMVMTAMWTGGDWLS 360
 Qy 361 NPEVDKMLSEVTNLIYHKNIPENAHVDFIWGLDAPHMYNEIHLMOEETNLSQGRCE 420
 Db 361 NPDVDKTLSEVTNLIYHKNIPENAHVDFIWGLDAPHMYNEIHLMOEETNLSQGRCE 419
 Qy 421 AVL 423
 Db 420 VKL 422

RESULT 4
 Q8C2G7 PRELIMINARY; PRT; 397 AA.
 AC Q8C2G7
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-WAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-WAR-2003 (Tremblrel. 23, Last annotation update)
 DE Lysosomal acid lipase 1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NOD; TISSUE=Thymus;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs."
 RL Nature 420:563-573 (2002).
 DR EMBL; AK088659; BAC0484.1; --
 SQ SEQUENCE 397 AA; 45325 MW; 866B123A1A058134 CRC64;

Query Match 54.4%; Score 1222.5; DB 11; Length 397;
 Best Local Similarity 56.7%; Pred. No. 2.4e-103;
 Matches 221; Conservative 70; Mismatches 96; Indels 3; Gaps 1;
 Qy 23 VAYMFQNVNVMFT---KAYDPEAFNMISIIHQGVPCPEEYEVATEDGYLSVRIP 79
 Db 6 LVFVFTIGILSRVPTGTVSAVDPEVNMVTEIIMRWGYPGEHSVLGTGDIHRIIP 65
 Qy 80 RGLVOPKPTGSPVLLQHLGVGASNMISNLPNNLSGLFILADAGFDVWNGNSRGN 139
 Db 66 RGRKHFGKGPFPVYLQHLGLASNMVNTIDNSLGLFLLADAGFDVWNGNSRGN 125
 Qy 140 KHTLSIDODEFWAFSYDEMARPDLPAVINFLQKTGOEKLYYGVSGTGMGTAFSTM 199
 Db 126 KHTLSVSDQDEFWAFSYDEMAYKDLPAVINFLQKTGOEKLYYGVSGTGMGTAFSTM 185
 Qy 200 PELAKIKMYFALAPIATVYHAKSPGKFLPLPDMWIKGLFGKKEFLYQTRFLRLQVLYL 259
 Db 186 PELAKIKMYFALAPIATVYHAKSPGKFLPLPDMWIKGLFGKKEFLYQTRFLRLQVLYL 245
 Qy 260 CQGVILQDQCSNIMLLGGFNTNNMNSRASYAAHTLAGTSVQNLHWSQAVNGELA 319
 Db 246 CTHVIMELCANVFLLCGFNENKLNMRDVTYTHCPAGTSVQNLHWSQAVNGELA 305
 Qy 320 FQWSEKYNFYHNSQFPPSYNKNMRLPTALWSGGRDLADINDITLLIQPKLYVHK 379
 Db 306 FQWSEKYNFYHNSQFPPSYNKNMRLPTALWSGGRDLADINDITLLIQPKLYVHK 365
 Qy 380 NPEVAHVDIWLGLDAPHMYNEIHLMOQ 409
 Db 366 NPEVDHLDIWLGLDAPHMYNEIHLMKK 395

RESULT 5
 Q8D798 PRELIMINARY; PRT; 395 AA.
 AC Q8D798
 DT 01-JUN-2001 (Tremblrel. 17, Created)

DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 GN 2310051B21RIK protein.
 GN 2310051B21RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Tongue;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Famanaka I.,
 RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli P., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,
 RA Suzuki H., Toyo-Oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawasumi H., Kohseki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690 (2001).
 DR EMBL; AK009431; BAB26283.1; --
 DR HSP; P07098; IHLG.
 DR MGD; MGI:1914967; 2310051B21RIK.
 DR InterPro; IPR000073; A/B_hydrolase.
 DR InterPro; IPR000734; Lipase.
 DR InterPro; IPR000379; Ser_estr_site.
 DR Pfam; PF00561; abhydrolase; 1.
 DR PROSITE; PS00120; LIPASE_SER; 1.
 SQ SEQUENCE 395 AA; 44603 MW; D3FD8B6PEA671B3E CRC64;

Query Match 50.5%; Score 1134.5; DB 11; Length 395;
 Best Local Similarity 53.5%; Pred. No. 2.8e-95;
 Matches 213; Conservative 70; Mismatches 108; Indels 7; Gaps 3;
 Qy 17 MWLLILVAYMFQ-RNVNSVHMPTKAVDPEAFNMISIIHQGVPCPEEYEVATEDGYLSV 75
 Db 1 MWLLITVSVLSAFGGAGLFGKLGPKNPNANVNSQMITWYGPSEYEVATEDGYLV 60
 Qy 76 NRIPRGLVOPKPTGSPVLLQHLGVGASNMISNLPNNLSGLFILADAGFDVWNGNSRGN 135
 Db 61 YRIPYKGNSENIGKFPVYLQHLGLASNTMTNLPNNLSGLFILADAGFDVWNGNSRGN 120
 Qy 136 ANRKHKTLSIDODEFWAFSYDEMARPDLPAVINFLQKTGOEKLYYGVSGTGMGTAF 195
 Db 121 TWSRKNVYSPDSVEFWAFSYDEMAYKDLPAVINFLQKTGOEKLYYGVSGTGMGTAF 180
 Qy 196 FSTMPLEAKIKMYFALAPIATVYHAKSPGKFLPLPDMWIKGLFGKKEFL---YQTRFL 252
 Db 181 FSTNPALAKIKMYFALAPIATVYHAKSPGKFLPLPDMWIKGLFGKKEFLYQTRFL 240
 Qy 253 RQVIVLCQGVILQDQCSNIMLLGGFNTNNMNSRASYAAHTLAGTSVQNLHWSQAV 312
 Db 241 GTEV---CSRLLDLCLSNALIFCGFDKKNLNSRLDVLGHNPAAGTSTQDLFWAQLA 297
 Qy 313 NSGELRAFQWSEKYNFYHNSQFPPSYNKNMRLPTALWSGGRDLADINDITLLIQPKLYVHK 372
 Db 298 KSGKQAYWNGSPQLQNLHYNKTPPYVDVSMVTPVIAVWNGHDLADINDITLLIQPKLYVHK 357
 Qy 373 TNLIIYHKNIPENAHVDFIWGLDAPHMYNEIHLMOQ 410
 Db 358 PNLIIYHKNIPENAHVDFIWGLDAPHMYNEIHLMOQ 395

```

RESULT 6
Q9D6Q6 PRELIMINARY; PRT; 395 AA.
ID Q9D6Q6
AC Q9D6Q6
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Adult male tongue cDNA, RIKEN full-length enriched library,
DE clone:2310067K20, full insert sequence.
GN 2310051B21RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nihi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Brownstein M.J., Bult C., Fletcher C., Carninci P., de Bonaldo M.F.,
RA Gustincich S., Hill D., Hofmann M., Rodriguez I., Sakamoto N.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ringwald M., Hume D.A., Kamiya M., Lee N.H.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK010093; BAB26697.1; -.
DR HSSP; P07098; 1HLG.
DR MGD; MGI:1914967; 2310051B21RIK.
DR InterPro; IPR000073; A/b_hydrolase.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser_estrs_site.
DR Pfam; PF00561; abhydrolase_1.
DR PROSITE; PS00120; LIPASE_SER; 1.
SQ SEQUENCE 395 AA; 44623 MW; 245055F5E7FF07C9 CRC64;

Query Match 50.5%; Score 1134.5; DB 11; Length 395;
Best Local Similarity 53.8%; Pred. No. 2.8e-95;
Matches 214; Conservative 69; Mismatches 108; Indels 7; Gaps 3;

QY 17 MWLLILVAYMFQ-RNVNSVHMTKAVDPEAFNISEIIHQGYPCREVEVATEDGYILSV 75
DB 1 MWLLVTSVLSAFGGAHGLFGKLPKPNANNNVSMITYGYSEEVVTEDEGYILGV 60

QY 76 NRIPGLVQPKTGRSPVVLQHLGVGGASNWSLNNPNSLGFILADAGFVWNGNSRGN 135
DB 61 YRIPYKKNSENIGRPVAYLQHLGVASATNWTNLPNNSLAFLADAGYDVLGNSRGN 120

QY 136 AWSRKHKLISIDQDFWAFSDYEMARFPLPAVINFILOKTOGKTYVGYSGTMTGFIA 195
DB 121 TWSRKNVYSPDSVFWAFSDFEMAKYDLPATIDFIVQKTOGKHYVGHSGGTIGFIA 180

QY 196 FSTMPALAKIKRYPALAPIATVHKASPGTKFLLLPDMMIKGLFGKKEFL---YQTRFL 252
DB 181 FSTNPALAKIKRYPALAPVATVKYTSFPPKSLIPKFLKLVIFGNKMFPHNYLQDFL 240

QY 253 RQLVYLCGVILDOICSNIMLLGGFWNNWNSRASYAAHLAGTSVQNLHWSQAV 312
DB 241 GTEV---CSRELLDLCLCSNAULFCGDKKNLNSRFDVYLGHNPAGSTQDLFHHQA 297

RESULT 7
Q9D6P3 PRELIMINARY; PRT; 395 AA.
ID Q9D6P3
AC Q9D6P3
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Adult male tongue cDNA, RIKEN full-length enriched library,
DE clone:2310069P19, full insert sequence.
GN 2310051B21RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nihi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Brownstein M.J., Bult C., Fletcher C., Carninci P., de Bonaldo M.F.,
RA Gustincich S., Hill D., Hofmann M., Rodriguez I., Sakamoto N.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ringwald M., Hume D.A., Kamiya M., Lee N.H.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK010139; BAB26725.1; -.
DR HSSP; P07098; 1HLG.
DR MGD; MGI:1914967; 2310051B21RIK.
DR InterPro; IPR000073; A/b_hydrolase.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser_estrs_site.
DR Pfam; PF00561; abhydrolase_1.
DR PROSITE; PS00120; LIPASE_SER; 1.
SQ SEQUENCE 395 AA; 44651 MW; E48A73CCCBFD359D CRC64;

Query Match 50.5%; Score 1134.5; DB 11; Length 395;
Best Local Similarity 53.8%; Pred. No. 2.8e-95;
Matches 214; Conservative 69; Mismatches 108; Indels 7; Gaps 3;

QY 17 MWLLILVAYMFQ-RNVNSVHMTKAVDPEAFNISEIIHQGYPCREVEVATEDGYILSV 75
DB 1 MWLLVTSVLSAFGGAHGLFGKLPKPNANNNVSMITYGYSEEVVTEDEGYILGV 60

QY 76 NRIPGLVQPKTGRSPVVLQHLGVGGASNWSLNNPNSLGFILADAGFVWNGNSRGN 135
DB 61 YRIPYKKNSENIGRPVAYLQHLGVASATNWTNLPNNSLAFLADAGYDVLGNSRGN 120

QY 136 AWSRKHKLISIDQDFWAFSDYEMARFPLPAVINFILOKTOGKTYVGYSGTMTGFIA 195
DB 121 TWSRKNVYSPDSVFWAFSDFEMAKYDLPATIDFIVQKTOGKHYVGHSGGTIGFIA 180

```

```

QY 196 FSTMPBLAQIKMYFALAPALATVYKAKSPGTKEFLLPDMWIKGLFKGKEFL---YOTRFL 252
DB 181 FSTNPALAKKIRFRFALAPALATVYKTESPFKKISLIPKFLVIFGNKMFPHNYLDQFL 240
QY 253 ROLVIVYLCQVILDQICSNIMLLGGFNTNNMNSRASVYAAHTLAGTSVONILHWSQAV 312
DB 241 GTEV---CSRELLDLLCSNALFICGDFCKKNLNVSRFDVYLGHNPACTSTQDLFWAQLA 297
QY 313 NSGELRAFDMGSETKNLEKCNQPTPVRYRVRDMTPTAMWTGGQDWLNSPBDVKNMLSEV 372
DB 298 KSGKLQAYNWGSPLOMLHYNQKTPPYDYVSANTVPIAVWNGGHDLADPDQDVAMLLPKL 357
QY 373 TNLIVYHKNIPWAHVDVFIWGLDAPHEMYNEIHLMOQE 410
DB 358 PNLLYHKEILPYNHLDFIWADAPQEVYNEIVTMMAED 395

RESULT 8
Q9D6X0 PRELIMINARY; PRT; 395 AA.
AC Q9D6X0;
DT 01-JUN-2001 (TREMELrel. 17, Created)
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE 2310051B21Rik protein.
GN 2310051B21Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Maignan J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK009875; BAB26556.1; -
DR HSSP; P07098; IHLG.
DR MCD; MG1:1514967; 2310051B21Rik.
DR InterPro; IPR000073; A/b hydrolase.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser estersite.
DR Pfam; PF00561; abhydrolase_1.
DR PROSITE; PS00120; LIPASE SER; 1.
SQ SEQUENCE 395 AA; 44665 MW; 40CNA6B67859A8C5B CRC64;

Query Match 50.5%; Score 1134.5; DB 11; Length 395;
Best Local Similarity 53.5%; Pred. No. 2.8e-95;
Matches 213; Conservative 70; Mismatches 108; Indels 7; Gaps 3;

QY 17 MWLLILVAYMFO-RNVNSVHMTKAVDPEAFNISEITQHOGYPCPEEYEVATEDGYILSV 75
DB 1 MWLLILVTSLSAFGGAHFLGKLGPKNPANNVSMITTYGPPSEYEVATEDGYILGV 60

```

```

QY 76 NRIPRGLVQPKTGRSPVLLQHLVGGASWISNLPNNSLGFILADAGFDVWNGSRGN 135
DB 61 YRIPYKGKNSENTGKRPVAYLOHGLTASATNTNLPNNSLAFILADAGFDVWNGSRGN 120
QY 136 AWSRKHTKTSIDQDEFWAFSYDEMAFDFLPVAVINFILQKTGQEKIYVYVGSQQTMGFTA 195
DB 121 TWSRKNVYSPDSVETWAFSDFEMAKYDLPATIDFVQKTGQEKIHYVGHSGQTIGFTA 180
QY 196 FSTMPBLAQIKMYFALAPALATVYKAKSPGTKEFLLPDMWIKGLFKGKEFL---YOTRFL 252
DB 181 FSTNPALAKKIRFRFALAPALATVYKTESPFKKISLIPKFLVIFGNKMFPHNYLDQFL 240
QY 253 ROLVIVYLCQVILDQICSNIMLLGGFNTNNMNSRASVYAAHTLAGTSVONILHWSQAV 312
DB 241 GTEV---CSRELLDLLCSNALFICGDFCKKNLNVSRFDVYLGHNPACTSTQDLFWAQLA 297
QY 313 NSGELRAFDMGSETKNLEKCNQPTPVRYRVRDMTPTAMWTGGQDWLNSPBDVKNMLSEV 372
DB 298 KSGKLQAYNWGSPLOMLHYNQKTPPYDYVSANTVPIAVWNGGHDLADPDQDVAMLLPKL 357
QY 373 TNLIVYHKNIPWAHVDVFIWGLDAPHEMYNEIHLMOQE 410
DB 358 PNLLYHKEILPYNHLDFIWADAPQEVYNEIVTMMAED 395

RESULT 9
Q9C9P7 PRELIMINARY; PRT; 395 AA.
AC Q9C9P7;
DT 01-JUN-2001 (TREMELrel. 17, Created)
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE 2310051B21Rik protein.
GN 2310051B21Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Maignan J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK010236; BAB26787.1; -
DR EMBL; AK009413; BAB26272.1; -
DR EMBL; AK009428; BAB26280.1; -
DR EMBL; AK009459; BAB26300.1; -
DR EMBL; AK009473; BAB26312.1; -
DR EMBL; AK009474; BAB26313.1; -
DR EMBL; AK009479; BAB26316.1; -
DR EMBL; AK009523; BAB26338.1; -
DR EMBL; AK009525; BAB26339.1; -
DR EMBL; AK009546; BAB26352.1; -
DR EMBL; AK009571; BAB26368.1; -

```



```

DR EMBL; AK009573; BAB26370.1; -
DR EMBL; AK009729; BAB26466.1; -
DR EMBL; AK009773; BAB26495.1; -
DR EMBL; AK010019; BAB26647.1; -
DR EMBL; AK010035; BAB26656.1; -
DR EMBL; AK010058; BAB26673.1; -
DR EMBL; AK010061; BAB26675.1; -
DR EMBL; AK010124; BAB26715.1; -
DR EMBL; AK010125; BAB26716.1; -
DR HSPF; P07098; IHLG.
DR MGD; MGI:1914967; 2310051B21Rik.
DR InterPro; IPR000073; A/b hydrolase.
DR InterPro; IPR000734; Lipase.
DR Pfam; PF00561; abhydrolase_1_site.
DR PROSITE; PS00120; LIPASE SER; 1.
SQ SEQUENCE 395 AA; 44637 MW; D3F96B65EA671E34 CRC64;

Query Match 50.4%; Score 1133.5; DB 11; Length 395;
Best Local Similarity 53.5%; Pred. No. 3.4e-95;
Matches 213; Conservative 70; Mismatches 108; Indels 7; Gaps 3;

QY 17 MWLLILVAYMQ-RNVNSVHPTKAVDPEAPNISEIIHQGYPCSEYEVATEDGYILSV 75
DB 1 MWLLLVTSVLSAFGGAHGLFGKLGPKNPEANMNVSMITYGYPSEYEVATEDGYILV 60

QY 76 NRIPRGLVQPKTGRPVLLQHLGVGASWISNLPNSLGFILADAGFDVWNGSRGN 135
DB 61 YRIPYKKNSENIKRPVAYLQHLIASATWITNLPNSLAFILADAGYDVWNGSRGN 120

QY 136 AWSRKHTLSIDQDEFWAFSDENARFDLPVINFILQKQEKIYVYVSGQTTMGFIA 195
DB 121 TWSRKNVYSPDSVEFWAFSDENAKYDLPATIDFIVQKQEKIHYVGHSGQTTGIFA 180

QY 196 FSTWPELAKIKMYFALAPIATVHAKSPGKFKLLPDMIKGLFGKKEFL--YQREFL 252
DB 181 FSTNPALAKIKRFVAPVATVYTESPFKKISLIPKLLKLVIFGNKMFPHNYLDQFL 240

QY 253 RQLVLYLCQGVILDOI CSNIMLLGGFNTNNMNSRASVYAAHTLAGTSVQNLHWSQAV 312
DB 241 GTEV---CSRELLDLCSNALFICGFDKKNLVSRFGVILGHNPACTSTQDLFHWQA 297

QY 313 NSGELRAFDSGETKLEKCNQPTPVRYRVDMTVPTAMTGGODWLSNEDVQMLLSEV 372
DB 298 KSGKLQAYNWGSPQNMLHYNQKTPPYDYDSAMTVPVAVWNGGHDILADPDQVAMLLPKL 357

QY 373 TNLIVHKNIPWAHVDFIWLGDAPHRMYNEIHLMOQE 410
DB 358 PNLLYHKELIPYNHLDFIWDAPQEVYNEIVTWMAED 395

RESULT 10
Q9C9P8 PRELIMINARY; PRT; 395 AA.
AC Q9C9P8;
DT 01-JUN-2001 (TREMELrel. 17, Created)
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
DE 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE 2310051B21Rik protein.
GN 2310051B21Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K. I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

```

```

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Mateo Y., Nikaio I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hall D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,
RA Hayashizaki Y.;
RT "functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK010116; BAB26711.1; -
DR EMBL; AK009300; BAB26201.1; -
DR HSPF; P07098; IHLG.
DR MGD; MGI:1914967; 2310051B21Rik.
DR InterPro; IPR000073; A/b hydrolase.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser_estrs_site.
DR Pfam; PF00561; abhydrolase_1.
DR PROSITE; PS00120; LIPASE SER; 1.
SQ SEQUENCE 395 AA; 44579 MW; D2296865EA671D34 CRC64;

Query Match 50.4%; Score 1133.5; DB 11; Length 395;
Best Local Similarity 53.5%; Pred. No. 3.4e-95;
Matches 213; Conservative 70; Mismatches 108; Indels 7; Gaps 3;

QY 17 MWLLILVAYMQ-RNVNSVHPTKAVDPEAPNISEIIHQGYPCSEYEVATEDGYILSV 75
DB 1 MWLLLVTSVLSAFGGAHGLFGKLGPKNPEANMNVSMITYGYPSEYEVATEDGYILV 60

QY 76 NRIPRGLVQPKTGRPVLLQHLGVGASWISNLPNSLGFILADAGFDVWNGSRGN 135
DB 61 YRIPYKKNSENIKRPVAYLQHLIASATWITNLPNSLAFILADAGYDVWNGSRGN 120

QY 136 AWSRKHTLSIDQDEFWAFSDENARFDLPVINFILQKQEKIYVYVSGQTTMGFIA 195
DB 121 TWSRKNVYSPDSVEFWAFSDENAKYDLPATIDFIVQKQEKIHYVGHSGQTTGIFA 180

QY 196 FSTWPELAKIKMYFALAPIATVHAKSPGKFKLLPDMIKGLFGKKEFL--YQREFL 252
DB 181 FSTNPALAKIKRFVAPVATVYTESPFKKISLIPKLLKLVIFGNKMFPHNYLDQFL 240

QY 253 RQLVLYLCQGVILDOI CSNIMLLGGFNTNNMNSRASVYAAHTLAGTSVQNLHWSQAV 312
DB 241 GTEV---CSRELLDLCSNALFICGFDKKNLVSRFGVILGHNPACTSTQDLFHWQA 297

QY 313 NSGELRAFDSGETKLEKCNQPTPVRYRVDMTVPTAMTGGODWLSNEDVQMLLSEV 372
DB 298 KSGKLQAYNWGSPQNMLHYNQKTPPYDYDSAMTVPVAVWNGGHDILADPDQVAMLLPKL 357

QY 373 TNLIVHKNIPWAHVDFIWLGDAPHRMYNEIHLMOQE 410
DB 358 PNLLYHKELIPYNHLDFIWDAPQEVYNEIVTWMAED 395

RESULT 11
Q9D6L9 PRELIMINARY; PRT; 395 AA.
ID Q9D6L9;
AC Q9D6L9;
DT 01-JUN-2001 (TREMELrel. 17, Created)
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
DE 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Adult male tongue cDNA, RIKEN full-length enriched library,
DE clone:23100761613, full insert sequence.
GN 2310051B21Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

```

```

RN  SEQUENCE FROM N.A.
RP  STRAIN=C57BL/6J; TISSUE=Tongue;
RX  MEDLINE=21085660; PubMed=11217851;
RA  Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA  Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA  Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA  Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA  Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA  Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA  Kuehl P., Lewis S., Matsuo Y., Nikola I., Pesole G., Quackenbush J.,
RA  Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA  Sakai H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA  Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA  Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,
RA  Hayaishizaki Y.,
RT  "Functional annotation of a full-length mouse cDNA collection.";
RL  Nature 409:685-690(2001).
DR  EMBL; AK010203; BAB26766.1; -.
DR  HSP; P07098; 1HLG.
DR  MGD; MGI:1914967; 2310051B21Rik.
DR  InterPro; IPR000073; A/b_hydrolase.
DR  InterPro; IPR000734; Lipase.
DR  InterPro; IPR000379; Ser_estrs_site.
DR  Pfam; PF00561; abhydrolase; 1.
DR  PROSITE; PS00120; LIPASE_SER; 1.
SQ  SEQUENCE 395 AA; 44671 MW; B8936162510AA55C CRC64;

Query Match 50.4%; Score 1131.5; DB 11; Length 395;
Best Local Similarity 53.5%; Pred. No. 5.2e-95;
Matches 213; Conservative 69; Mismatches 109; Indels 7; Gaps 3;

QY 17 MWLLLVAYMFQ-RNNVSVHMPKAVDPPEAFNMISEIIHQGYPCBEYEVATEDGYILSV 75
DB 1 MWLLLVTSVLSAFGAGLFGKLGPKNPENNVNSQMITWGYPSBEYEVATEDGYILGV 60

QY 76 NRIPRGLVQPKKTSRPRVLLQHLVGGASNWTSLNPNNSLGFILADAGPDVWNGNSRG 135
DB 61 YRIPYKKNSENIGKRPVAYLQGLTASATNWTNLPNNLSLAFILADAGYDVLWNGNSRG 120

QY 136 AWSRKHKTLSIDQDEFWAFSDEWAFDLPVAVINFILQKTGOEKIYVGYSGQTWGFIA 195
DB 121 TWSRKNVYSPDSVEFWAFSDEWAFDLPVAVINFILQKTGOEKIYVGYSGQTWGFIA 180

QY 196 FSTMPELAOKIKMYFALAPATVTKHAKSPGTEKLLPDMKIKGLFGKKEPL---YQTRFL 252
DB 181 FSTNPALAKKIKRFPYALAPVATVKYTESPFKKISLPKFLKVIKGNKMFPHNYLDQFL 240

QY 253 RQLVYLCCGVILDOICSNIMLLGGFNTNMMNSRASVYAHTLAGTSVQNLHWSQAV 312
DB 241 GTEV---CSRELLDLCSNALFFCGFDKQLNVSRPDVYLGHPAGTQDLFWHAQLA 297

QY 313 NSGELRAPDWSGSETKNLEKCNQPTPVRYVRDVTPTAMTGGDWLSNPDVKMLLSEV 372
DB 298 KSGKLOAYNWGSPLOKMLHYNQKSPYDYVSATVPIAVWNGGHDILADQDVAMLLPKL 357

QY 373 TNLVYHKNIPENAHVDFIWLGDAPHMYNEIHLMOQE 410
DB 358 PNLLYHKEILPYNLHDFIWMADAPQEVYNEIVTMMAED 395

RESULT 12
Q9D6T5 ID Q9D6T5 PRELIMINARY; PRT; 395 AA.
AC Q9D6T5
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)

```

```

DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Adult male tongue cDNA, RIKEN full-length enriched library,
GN clone:2310061A13, full insert sequence.
GN 2310051B21R1K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikola I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Bono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,
RA Hayaishizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK009990; BAB26629.1; -.
DR HSP; P07098; 1HLG.
DR MGD; MGI:1914967; 2310051B21Rik.
DR InterPro; IPR000073; A/b_hydrolase.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser_estrs_site.
DR Pfam; PF00561; abhydrolase; 1.
DR PROSITE; PS00120; LIPASE_SER; 1.
SQ SEQUENCE 395 AA; 44623 MW; D3F96DB83161C3BF CRC64;

```

```

Query Match 50.3%; Score 1129.5; DB 11; Length 395;
Best Local Similarity 53.3%; Pred. No. 7.9e-95;
Matches 212; Conservative 71; Mismatches 108; Indels 7; Gaps 3;

QY 17 MWLLLVAYMFQ-RNNVSVHMPKAVDPPEAFNMISEIIHQGYPCBEYEVATEDGYILSV 75
DB 1 MWLLLVTSVLSAFGAGLFGKLGPKNPENNVNSQMITWGYPSBEYEVATEDGYILGV 60

QY 76 NRIPRGLVQPKKTSRPRVLLQHLVGGASNWTSLNPNNSLGFILADAGPDVWNGNSRG 135
DB 61 YRIPYKKNSENIGKRPVAYLQGLTASATNWTNLPNNLSLAFILADAGYDVLWNGNSRG 120

QY 136 AWSRKHKTLSIDQDEFWAFSDEWAFDLPVAVINFILQKTGOEKIYVGYSGQTWGFIA 195
DB 121 TWSRKNVYSPDSVEFWAFSDEWAFDLPVAVINFILQKTGOEKIYVGYSGQTWGFIA 180

QY 196 FSTMPELAOKIKMYFALAPATVTKHAKSPGTEKLLPDMKIKGLFGKKEPL---YQTRFL 252
DB 181 FSTNPALAKKIKRFPYALAPVATVKYTESPFKKISLPKFLKVIKGNKMFPHNYLDQFL 240

QY 253 RQLVYLCCGVILDOICSNIMLLGGFNTNMMNSRASVYAHTLAGTSVQNLHWSQAV 312
DB 241 GTEV---CSRELLDLCSNALFFCGFDKQLNVSRPDVYLGHPAGTQDLFWHAQLA 297

QY 313 NSGELRAPDWSGSETKNLEKCNQPTPVRYVRDVTPTAMTGGDWLSNPDVKMLLSEV 372
DB 298 KSGKLOAYNWGSPLOKMLHYNQKSPYDYVSATVPIAVWNGGHDILADQDVAMLLPKL 357

QY 373 TNLVYHKNIPENAHVDFIWLGDAPHMYNEIHLMOQE 410
DB 358 PNLLYHKEILPYNLHDFIWMADAPQEVYNEIVTMMAED 395

```

```
RESULT 13
Q9D767
ID AC Q9D767 PRELIMINARY; PRT; 395 AA.
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE 2310051B21Rik protein.
GN 2310051B21Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stauber F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyshew-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL; AK009537; BAB26346.1; -.
DR HSSP; P07098; 1HLG.
DR MGD; MGI:1914967; 2310051B21Rik.
DR InterPro; IPR000073; A/b hydrolase.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser_estr_site.
DR Pfam; PF00561; abhydrolase; 1.
DR PROSITE; PS00120; LIPASE SER; 1.
SQ SEQUENCE 395 AA; 44607 MW; E43317C2254FA8FB CRC64;

Query Match 50.2%; Score 1128.5; DB 11; Length 395;
Best Local Similarity 53.3%; Pred. No. 9.8e-95;
Matches 212; Conservative 70; Mismatches 109; Indels 7; Gaps 3;

QY 17 MWLLILVAFMQ-RNVNSVHPTKAVDPEAFMNISEIIHQGYPCPEEYEVATEDGYILSV 75
DB 1 MWLLVTSVLSAFGAGHGLGKLGPKNPEANMVSQMTYWGYPSEEVVTEGYILGV 60

QY 76 NRIPGLVQPKTSRPPVLLQHLGVGGASWISNLPNNSLGFILADAGFDVWNGSRGN 135
DB 61 YRIPYKKNSENIGKRPVAYLQHLGASATNITLNPNSLAFILADAGYDVLNLSRGN 120

QY 136 AMSRKHKTLISIDQDFWAFSPYDEMARFOLPAVINFILOKTQEKIYVYGSQGTMGFIA 195
DB 121 TWSRKNVYSPDSVEFWAFSPDEMAKYDLPATIDFIVQKTQEKIHYVGHSGTIGFIA 180

QY 196 FSTMPELLAQIKMYFALAPIATVTKHAKSPGTFKLLPDMMKLPGKKEFL---YQTRFL 252
DB 241 GTEV---CSRLLDILCSNALFIFCGFDKKNLNSRFDVYLGHNPAGTSTQDLFFHQA 297
```

```
QY 313 NSGELRAFDWGSETKXLEKCKNPTVRYRVEDMTVPTAMWTGGQDMLSNPDEDVKMLLSEV 372
DB 298 KSGKLOAYNWGSEFLQNMLHYNQKTPPYDYDSAMTVPTAVWNGGHDILADPDQVALLPKL 357

QY 373 TNLIIYHKNPWAHVDFIWLGLDAPHRMYNEIHLMOQE 410
DB 358 PNLLYHKEILPYNHLDFIWMADPOEVNEIWTWMAED 395

RESULT 14
Q9D7C5
ID AC Q9D7C5 PRELIMINARY; PRT; 395 AA.
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE 2310051B21Rik protein.
GN 2310051B21Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stauber F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyshew-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL; AK009359; BAB26240.1; -.
DR HSSP; P07098; 1HLG.
DR MGD; MGI:1914967; 2310051B21Rik.
DR InterPro; IPR000073; A/b hydrolase.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser_estr_site.
DR Pfam; PF00561; abhydrolase; 1.
DR PROSITE; PS00120; LIPASE SER; 1.
SQ SEQUENCE 395 AA; 44748 MW; 3ED43992458DE058 CRC64;

Query Match 50.2%; Score 1127.5; DB 11; Length 395;
Best Local Similarity 53.3%; Pred. No. 1.2e-94;
Matches 212; Conservative 69; Mismatches 110; Indels 7; Gaps 3;

QY 17 MWLLILVAFMQ-RNVNSVHPTKAVDPEAFMNISEIIHQGYPCPEEYEVATEDGYILSV 75
DB 1 MWLLVTSVLSAFGAGHGLGKLGPKNPEANMVSQMTYWGYPSEEVVTEGYILGV 60

QY 76 NRIPGLVQPKTSRPPVLLQHLGVGGASWISNLPNNSLGFILADAGFDVWNGSRGN 135
DB 61 YRIPYKKNSENIGKRPVAYLQHLGASATNITLNPNSLAFILADAGYDVLNLSRGN 120

QY 136 AMSRKHKTLISIDQDFWAFSPYDEMARFOLPAVINFILOKTQEKIYVYGSQGTMGFIA 195
DB 121 TWSRKNVYSPDSVEFWAFSPDEMAKYDLPATIDFIVQKTQEKIHYVGHSGTIGFIA 180

QY 196 FSTMPELLAQIKMYFALAPIATVTKHAKSPGTFKLLPDMMKLPGKKEFL---YQTRFL 252
```

```

Db 181 FSTPALAKKIKRYALAPVATVTKYSPFKKHIFPKFLKVIKFGNKMFMHNYLDQFL 240
QY 253 RQLVYLCOGQVILDOICSNIMLLGGFNTNNMNSRASVYAAHTLAGTSVQNILHWSQAV 312
Db 241 GTEV---CSRELLDLCSNLFIFCGFDKKNLNVSRFDVYLGHNPAGTSTQDLFHWQA 297
QY 313 NSGELRPFADWGSSETKNLEKCNQPTPVRYRVRDMTPTAMTGGQDWLSNPDVVKMLSEV 372
Db 298 KSGKLQAYNWGSPQLNMLHYNQKTPPYDYVSAMTPTAVWNGGHDILADPDQVAMLLPKL 357
QY 373 TNLVYHKNIPEWAHVDFIWMGLDAPHRMNYEIIHLMQOE 410
Db 358 PNLLYHKEILPYNHLDFIWMADAPQEVNYEIVTMAED 395

RESULT 15
Q9D6L1 PRELIMINARY; PRT; 395 AA.
AC Q9D6L1
ID Q9D6L1
AD Q9D6L1
DT 01-JUN-2001 (TREMELrel. 17, Created)
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Adult male tongue cDNA, RIKEN full-length enriched library,
DE clone:2310079Q20, full insert sequence.
GN 2310051B21RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itch M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K. I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Pleischmann W., Gaasterland T., Gies C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK010231; BAB26784.1; --
DR ZSP; P07098; iHLG.
DR MGD; MGI:1914967; 2310051B21RIK.
DR InterPro; IPR000073; A/b hydrolase.
DR InterPro; IPR000734; Lipase.
DR InterPro; IPR000379; Ser esterase site.
DR Pfam; PF00561; abhydrolase; 1_
DR PROSITE; PS00120; LIPASE SER; 1.
SQ SEQUENCE 395 AA; 44659 MW; 13E0BB95FC5A40E0 CRC64;

```

```

Query Match 50.1%; Score 1126.5; DB 11; Length 395;
Best Local Similarity 53.5%; Pred. No. 1.5e-94;
Matches 213; Conservative 69; Mismatches 109; Indels 7; Gaps 3;

QY 17 MWLLILVAFQ-RNVNSVHMTKAVDPEAFNISEIICHQYPCPEEYEVATEDGYILSV 75
Db 1 MWLLVTSVLSAFGGAGLFGKLGPKNPNPANNVSMQITWYGPSEYEVVTEGYILGV 60
QY 76 NRIPRGLVQPKTGRPVLLQHLVGGASWISNLPNNSLGFIADAGFDVWNGNSRGN 135

```

```

Db 61 YRIPYKKNSENIGRRPVAYIQHGLIASATNITNLPNNSLAFILADAGYDVWLGNSRGN 120
QY 136 AWSRKHTLSIDQDEFWAFSPYDEMARFDLPVAVINFILQKTQOEKIYYVYSGQTTMGFIA 195
Db 121 TWSRKNVYSPDSVEFWAFSPDEMAKYDLPATIDFVQKTQOEKIHYVGHSGQTTIGFIA 180
QY 196 FSTMPDLAQKIKMYPALAPIATVKHAKSPGTFKLLPDMMIKGLPGKKEFL---YQTRFL 252
Db 181 FSTNPALAKKIKRFFVALAPVATVKYTESPFKISLIHKFLKVIKFGNKMFMHNYLDQFL 240
QY 253 ROLVYLCQGVILDOICSNIMLLGGFNTNNMNSRASVYAAHTLAGTSVQNILHWSQAV 312
Db 241 GTEV---CSRELLDLCSNLFIFCGFDKKNLNVSRFDVYLGHNPAGTSTQDLFHWQA 297
QY 313 NSGELRPFADWGSSETKNLEKCNQPTPVRYRVRDMTPTAMTGGQDWLSNPDVVKMLSEV 372
Db 298 KSGKLQAYNWGSPQLNMLHYNQKTPPYDYVSAMTPTAVWNGGHDILADPDQVAMLLPKL 357
QY 373 TNLVYHKNIPEWAHVDFIWMGLDAPHRMNYEIIHLMQOE 410
Db 358 PNLLYHKEILPYNHLDFIWMADAPQEVNYEIVTMAED 395

```

Search completed: February 19, 2004, 14:22:24
Job time : 35.1771 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 04:41:57 ; Search time 7514.62 Seconds
(without alignments)
11127.553 Million cell updates/sec

Title: US-10-042-431-45
Perfect score: 2044
Sequence: 1 gtcgaccacgcgtccggg.....aaaaaaagggcgccgc 2044

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:**

- 1: gb_ba:**
- 2: gb_htg:**
- 3: gb_in:**
- 4: gb_on:**
- 5: gb_ov:**
- 6: gb_pat:**
- 7: gb_ph:**
- 8: gb_pl:**
- 9: gb_pr:**
- 10: gb_ro:**
- 11: gb_sts:**
- 12: gb_sy:**
- 13: gb_un:**
- 14: gb_vi:**
- 15: em_ba:**
- 16: em_fun:**
- 17: em_hum:**
- 18: em_in:**
- 19: em_mu:**
- 20: em_on:**
- 21: em_or:**
- 22: em_ov:**
- 23: em_pat:**
- 24: em_ph:**
- 25: em_pl:**
- 26: em_ro:**
- 27: em_sts:**
- 28: em_un:**
- 29: em_vi:**
- 30: em_htg_hum:**
- 31: em_htg_inv:**
- 32: em_htg_other:**
- 33: em_htg_mus:**
- 34: em_htg_pin:**
- 35: em_htg_rod:**
- 36: em_htg_mam:**
- 37: em_htg_vrt:**
- 38: em_sy:**
- 39: em_htgo_hum:**
- 40: em_htgo_mus:**
- 41: em_htgo_other:**

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1138	55.7	1210	6	AX521798 Sequence
2	1107.8	54.2	1272	6	AX477670 Sequence
3	1000.4	48.9	2577	10	BC031933 Mus muscu
					BC031933 Human DNA
5	892	43.6	43456	9	AL353113 Homo sapi
					AC015461 Homo sapi
6	890.4	43.6	158195	2	AC015461 Homo sapi
					AX521801 Sequence
7	752	36.8	1116	6	AX521801 Sequence
					AC015506 Homo sapi
8	630.2	30.8	76553	2	AC015506 Homo sapi
					AX091558 Homo sapi
9	528.2	25.8	2982	9	AX091558 Homo sapi
					AX411234 Sequence
10	527.8	25.8	2481	6	AX411234 Sequence
					HSLYACLY
11	527.8	25.8	2481	9	HSU08464
					U08464 Human lysos
12	527.8	25.8	2626	9	HSLAL
					X76488 H. sapiens m
13	526.2	25.7	2493	6	BD094072
					BD094072 Shear str
14	526.2	25.7	2493	9	HUMLIPCHL
					BC012287 Homo sapi
15	524.6	25.7	2586	9	BC012287
					A26689 Precursor o
16	489.2	23.9	1378	6	A26689
					A26690 Precursor o
17	489.2	23.9	1378	6	A26690
					X05997 Human mRNA
18	470	23.0	1365	9	HSGLR
					A01046 H. sapiens m
19	470	23.0	1367	6	A01046
					A12714 Pregastric
20	470	23.0	1367	6	A12714
					A01157 R.norvegicu
21	467.6	22.9	1336	6	A01157
					X02309 Rat mRNA fo
22	467.6	22.9	1355	10	RNLIP
					AK096406 Homo sapi
23	462.2	22.6	2506	9	AK096406
					Z31689 M.musculus
24	461	22.6	2358	10	MMLYACLY
					S81497 lysosomal a
25	459.8	22.5	3144	10	S81497
					L26319 Bos taurus
26	458	22.4	1417	4	BOVME
					AX392821 Sequence
27	454	22.2	1228	6	AX392821
					AX392824 Sequence
28	448.4	21.9	1473	6	AX392824
					AX472752 Sequence
29	446.8	21.9	1185	6	AX472752
					AR039022 Sequence
30	444.2	21.7	1137	6	AR039022
					AR039023 Sequence
31	444.2	21.7	1137	6	AR039023
					AR039023 Sequence
32	444.2	21.7	1140	6	AR039023
					AR092634 Sequence
33	444.2	21.7	1140	6	AR092634
					AR092634 Sequence
34	444.2	21.7	1146	6	AR092634
					AR092635 Sequence
35	444.2	21.7	1146	6	AR092635
					AS7760 Sequence 5
36	444.2	21.7	1198	6	AS7760
					AS7756 Sequence 1
37	444.2	21.7	1528	6	AS7756
					AR039021 Sequence
38	444.2	21.7	1528	6	AR039021
					AR092632 Sequence
39	444.2	21.7	1528	6	AR092632
					Y13899 Canis fami
40	444.2	21.7	1651	4	CELIIPASE
					BC052131 Danio rer
41	441.8	21.6	1800	5	BC052131
					A39301 Sequence 2
42	441	21.6	1137	6	A39301
					A39303 Sequence 4
43	441	21.6	1140	6	A39303
					A39305 Sequence 6
44	441	21.6	1146	6	A39305
					A39300 Sequence 1
45	441	21.6	1528	6	A39300

ALIGNMENTS

RESULT 1
AX521798
LOCUS AX521798 1210 bp DNA linear PAT 05-OCT-2002
DEFINITION Sequence 1 from Patent WO0236731.
ACCESSION AX521798
VERSION AX521798.1 GI:23572854
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Xiao Y.
TITLE Human lysosomal acid lipase
JOURNAL Patent: WO 0236731-A 1 10-MAY-2002;

```

Bayer Aktiengesellschaft (DE)
Location/Qualifiers
1..1210
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
CDS
1..1197
/feature="unnamed protein product"
/codon_start=1
/protein_id="CAD52970.1"
/db_xref="GI:23572855"
/translation="MELKLPREDVLRWAEPTPLDRKQSEIIHQHQPCEBEVATED
GYLSVNPRIPLGLVOPKLTGSPVVLVHGLVGGASWISLNNPNSIKFTILADAGDV
WNGRGNWAKSRKHTLISIDQEPWAFSYDEWAPDLPAVINFILOKGTGQEKIYVGY
SQGTMGFASTMEPLACKIYFALAPIATVPHAKSGPGKFLFLLDMIKGLFGRK
EFLYQTRFLQVILVLCQGVILDDQICSNIMLLGGFTNNMNSRASVYAAHTLUGRS
VQNIHLWQAVNSGELRAFQWSETKNLEKNQPTPVRYRVRDNTVPTAMWTGGODWL
SNPEDIWMLLSVLLIYHKNIPWAHVDFIWLDPAPRMYNEIHLMQQBETNLSQ
RCEAVL"
BASE COUNT 338 a 264 c 296 g 312 t
ORIGIN
Query Match 55.7%; Score 1138; DB 6; Length 1210;
Best Local Similarity 100.0%; Pred. No. 5.5e-281;
Matches 1138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 273 AGTGAATCATCAACATCAAGGCTATCCCTGTGAGGAATATGAAGTCGCAACTGAAGAT 332
DB 73 AGTGAATCATCAACATCAAGGCTATCCCTGTGAGGAATATGAAGTCGCAACTGAAGAT 132
QY 333 GGGTATATCTTCTGTTTAAACAGGATTCCTCGAGGCTAGTGCAACCTAAGAAGACAGGT 392
DB 133 GGGTATATCTTCTGTTTAAACAGGATTCCTCGAGGCTAGTGCAACCTAAGAAGACAGGT 192
QY 393 TCCAGGCTGTGGTGTATCTGAGCATGCGCTAGTTGGAGGCTAGCAACTGGATTCC 452
DB 193 TCCAGGCTGTGGTGTATCTGAGCATGCGCTAGTTGGAGGCTAGCAACTGGATTCC 252
QY 453 AACCTGCCCAACAATAGCTGGGCTTCATTCTGGCAGATGCTGGTTTGGCTGTGGATG 512
DB 253 AACCTGCCCAACAATAGCTGGGCTTCATTCTGGCAGATGCTGGTTTGGCTGTGGATG 312
QY 513 GGGAAACAGCAGGGGAACGCTGGTCTGAAAACACAGACACTCTCCATAGACCAAGAT 572
DB 313 GGGAAACAGCAGGGGAACGCTGGTCTGAAAACACAGACACTCTCCATAGACCAAGAT 372
QY 573 GAGTCTCGGCTTTCAGTTATGATGAGTGGCTAGTGTTCAGCTTCCTCAGTGTATAAC 632
DB 373 GAGTCTCGGCTTTCAGTTATGATGAGTGGCTAGTGTTCAGCTTCCTCAGTGTATAAC 432
QY 633 TTTATTTTCAGAAAACGGGCCAGGAAAGATCTATATGTGCGGTATTCACAGGGCACC 692
DB 433 TTTATTTTCAGAAAACGGGCCAGGAAAGATCTATATGTGCGGTATTCACAGGGCACC 492
QY 693 ACCATGGCTTTATTCAGTTTCCACATGCCAGAGCTGGCTCAGAAAATCAAAATGTAT 752
DB 493 ACCATGGCTTTATTCAGTTTCCACATGCCAGAGCTGGCTCAGAAAATCAAAATGTAT 552
QY 753 TTTGCTTTAGCAACCCATAGCCATGTTTAAGCATGCAAAAAGCCCGGACCAAAATTTTG 812
DB 553 TTTGCTTTAGCAACCCATAGCCATGTTTAAGCATGCAAAAAGCCCGGACCAAAATTTTG 612
QY 813 TTGCTGCCAGATATGATGATCAAGGATTTGTTGGCAAAAAGAAATTTCTGTATCAGACC 872
DB 613 TTGCTGCCAGATATGATGATCAAGGATTTGTTGGCAAAAAGAAATTTCTGTATCAGACC 672
QY 873 AGATTTCTCAGACAACTGTTTATTTACCTTTGTGGCCAGGTGATTTCTGATCAGATTTGT 932
DB 673 AGATTTCTCAGACAACTGTTTATTTACCTTTGTGGCCAGGTGATTTCTGATCAGATTTGT 732
QY 933 AGTAAATATCATTTACTTCTGGGTGGATTCACACCAACATATGACATGACCCGAGCA 992
DB 733 AGTAAATATCATTTACTTCTGGGTGGATTCACACCAACATATGACATGACCCGAGCA 792
QY 993 AGTGTATATGCTGCCCAACAACCTCTTGTGGAACATCTGTGCAAAATATTTCTACACTGGAGC 1052
DB 793 AGTGTATATGCTGCCCAACAACCTCTTGTGGAACATCTGTGCAAAATATTTCTACACTGGAGC 852
QY 1053 CAGGCAGTGAATTTCTGTGAACTCCGGGCATTTGACTGGGGAGTGAGACCAAAATCTG 1112
DB 853 CAGGCAGTGAATTTCTGTGAACTCCGGGCATTTGACTGGGGAGTGAGACCAAAATCTG 912
QY 1113 GAAAAATGCAATCAGCAACCTCTCTGTAGGTACAGAGTCAGAGATGACGGTCCCTACA 1172
DB 913 GAAAAATGCAATCAGCAACCTCTCTGTAGGTACAGAGTCAGAGATGACGGTCCCTACA 972
QY 1173 GCAATGTGACAGAGGCTCAGACTGGCTTCAAAATCCAGAGAGCTGAAAAATGCTGCTC 1232
DB 973 GCAATGTGACAGAGGCTCAGACTGGCTTCAAAATCCAGAGAGCTGAAAAATGCTGCTC 1032
QY 1233 TCTGAGGTGACCAACCTCTCTCAATAAGAAATATTTCTGAATGGGCTCAGTGGATTTC 1292
DB 1033 TCTGAGGTGACCAACCTCTCTCAATAAGAAATATTTCTGAATGGGCTCAGTGGATTTC 1092
QY 1293 ATCTGGGTTTGGATGCTCTCTCAGCTATGTAATGAATCATCATCTGATGACAGCAG 1352
DB 1093 ATCTGGGTTTGGATGCTCTCTCAGCTATGTAATGAATCATCATCTGATGACAGCAG 1152
QY 1353 GAGGAGACCAACCTTTCCAGGAGCAGCTGTGAGGCGGTATTTGGAAGCATCTGACACT 1410
DB 1153 GAGGAGACCAACCTTTCCAGGAGCAGCTGTGAGGCGGTATTTGGAAGCATCTGACACT 1210
RESULT 2
AX477670 1272 bp DNA linear PAT 12-AUG-2002
LOCUS Sequence 8 from Patent WO0246418.
DEFINITION AX477670
ACCESSION AX477670
VERSION AX477670.1 GI:22216841
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 Griffin J.A., Gandhi A.R., Ramkumar J., Tang Y.T., Ding L., Yue H.,
AUTHORS Gietzen K.J., Sapperstein S.K., Honchell C.D., Bruns C.M.,
Duggan B.M., Xu Y. and Lee S.
TITLE Lipid-associated molecules
JOURNAL Patent: WO 0246418-A 8 13-JUN-2002;
Incyte Genomics, Inc. (US)
FEATURES
Location/Qualifiers
1..1272
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/notes="Incyte ID No: 7483978CB1"
BASE COUNT 351 a 263 c 311 g 347 t
ORIGIN
Query Match 54.2%; Score 1107.8; DB 6; Length 1272;
Best Local Similarity 97.2%; Pred. No. 3.1e-273; Indels 21; Gaps 1;
Matches 1146; Conservative 0; Mismatches 12;
QY 240 AAAGCTGTGGACCCAGAGCAATTCATGAATATTAGTGAATCATCAACATCAAGGCTAT 299
DB 94 AATGAATGAATCTGAGGTGCGATGAATAGTGAATCATCAACATCAAGGCTAT 153
QY 300 CCCTGTGAGGAATATGAAGTCGCAACTGAAGATGGGTATATCTTTCTGTATACAGATT 359
DB 154 CCCTGTGAGGAATATGAAGTCGCAACTGAAGATGGGTATATCTTTCTGTATACAGATT 213
QY 360 CCTCGAGGCTTAGTGCAACCTTAAGAAGACAGGTTCCAGGCTGTGGTGTACTGCAGCAT 419
DB 214 CCTCGAGGCTTAGTGCAACCTTAAGAAGACAGGTTCCAGGCTGTGGTGTACTGCAGCAT 273

```

QY 420 GGCTAGTTCGAGTGTCTAGCACTGATTTCCAACTGCTGCCCAACATAGCTGGCTTC 479
 Db 274 GGCTAGTTCGAGTGTCTAGCACTGATTTCCAACTGCTGCCCAACATAGCTGGCTTC 333
 QY 480 ATTCTGCAGATGCTGTTTGGATGCTGGATGGGAACAGCGGGAACGCTGCTCT 539
 Db 334 ATTCTGCAGATGCTGTTTGGATGCTGGATGGGAACAGCGGGAACGCTGCTCT 393
 QY 540 CGAAACACAGACACTCTCCATAGACCAAGATGCTTCTGGCTTTCAGTTATGATGAG 599
 Db 394 CGAAACACAGACACTCTCCATAGACCAAGATGCTTCTGGCTTTCAGTTATGATGAG 453
 QY 600 ATGCTAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
 Db 454 ATGCTAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 513
 QY 660 AAGATCTATTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 719
 Db 514 AAGATCTATTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 573
 QY 720 ATGCCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 779
 Db 574 ATGCCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 633
 QY 780 AAGCATGCAAAAGCCCGGGAACCAATTTTGTGCTGCTGCTGCTGCTGCTGCTGCT 839
 Db 634 AAGCATGCAAAAGCCCGGGAACCAATTTTGTGCTGCTGCTGCTGCTGCTGCTGCT 693
 QY 840 TTGTTTGGCAAAAGAAATTTCTGATCAGACCAAGATTTCTGATCAGACCAAGATTT 899
 Db 594 TTGTTTGGCAAAAGAAATTTCTGATCAGACCAAGATTTCTGATCAGACCAAGATTT 753
 QY 900 CTTTGTGCGCAGGTGATTTCTGATCAGATTTGATGATGATGATGATGATGATGATG 959
 Db 754 CTTTGTGCGCAGGTGATTTCTGATCAGATTTGATGATGATGATGATGATGATGATG 813
 QY 960 TTCAACCAACCAATATGACAT-----GAGCCGACGCAAGTGTA 998
 Db 814 TTCAACCAACCAATATGACATGTAATGTAATGTAATGTAATGTAATGTAATGTA 873
 QY 999 TATGTCGCCACACTTTGCTGGAACATCTGTGCAAAATATTTTACATGCTGAGCCAGGCA 1058
 Db 874 TATGTCGCCACACTTTGCTGGAACATCTGTGCAAAATATTTTACATGCTGAGCCAGGCA 933
 QY 1059 GTGAATTCGTGTAACCTCCGGGCAATTTGCTGCGGGAGTGAGACCAAAATCTGGAATA 1118
 Db 934 GTGAATTCGTGTAACCTCCGGGCAATTTGCTGCGGGAGTGAGACCAAAATCTGGAATA 993
 QY 1119 TGCAATCAGCCAACTCCTGTAAGGTACAGAGTACAGATATGACGGTCCCTACAGCAATG 1178
 Db 994 TGCAATCAGCCAACTCCTGTAAGGTACAGAGTACAGATATGACGGTCCCTACAGCAATG 1053
 QY 1179 TGGACAGAGTACAGAGTACAGAGTACAGAGTACAGAGTACAGAGTACAGAGTACAGAG 1238
 Db 1054 TGGACAGAGTACAGAGTACAGAGTACAGAGTACAGAGTACAGAGTACAGAGTACAGAG 1113
 QY 1239 GTGACCAACCTCATCTACCAAGAAATTTCTGCAATGCGGTCTGCGGGATTTTCTGCTGG 1298
 Db 1114 GTGACCAACCTCATCTACCAAGAAATTTCTGCAATGCGGTCTGCGGGATTTTCTGCTGG 1173
 QY 1299 GGTTCGATGCTCCTACCGTATGTACAAATGAAATCATTCATTCATTCATTCATTCATTC 1358
 Db 1174 GGTTCGATGCTCCTACCGTATGTACAAATGAAATCATTCATTCATTCATTCATTCATTC 1233
 QY 1359 ACCAACCTTTCCAGGAGCGGTGTGAGCGCGTATTGTGA 1397
 Db 1234 ACCAACCTTTCCAGGAGCGGTGTGAGCGCGTATTGTGA 1272

RESULT 3
 BC031933
 LOCUS
 DEFINITION Mus musculus, RIKEN cdna 2577 bp mRNA linear ROD 20-SEP-2002
 2577 bp mRNA linear ROD 20-SEP-2002

ACCESSION VERSION KEYWORDS SOURCE ORGANISM

REFERENCE AUTHORS TITLE JOURNAL

REMARK COMMENT

FEATURES source

CDS

BASE COUNT ORIGIN

Query Match 48.9%; Score 1000.4; DB 10; Length 2577;
 Best Local Similarity 83.8%; Pred. No. 1.2e-245;
 Matches 1170; Conservative 0; Mismatches 216; Indels 10; Gaps 3;
 QY 48 TTTTAAACCCCAAAATCTTCTTACTTAGAATAGTTGTATGATGCGCAGGAA-AAAA 106
 Db 33 TCTTCATCTCTGTTCTTCTTCTGACTCTAGATGATGTTGTATCATCTCTAGCAACAAAG 92

IMAGE:4020396, mRNA, complete cds.
 BC031933
 BC031933.1 GI:21594465
 MGC.
 Mus musculus (house mouse)
 Mus musculus
 Mammalia; Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Eukaryota; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 1 (bases 1 to 2577)
 Strausberg, R.
 Direct Submissions
 Submitted (06-JUN-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: http://mgc.nci.nih.gov
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: http://www.hgsc.bcm.tmc.edu/cdna/
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
 Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
 Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov
 Series: IRAC Plate: 30 Row: K Column: 3
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis, Similarity but not identity to protein.

Location/Qualifiers

1..2577
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="CZECH II"
 /db_xref="taxon:10090"
 /clone="MGC:25790 IMAGE:4020396"
 /tissue_type="Mammary tumor metastasized to lung."
 MMTV-LTR/Mnt1 model. Expression driven by an MMTV-LTR
 enhancer."
 /clone_lib="NCI CGAP_Lu30"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 112..1380
 /codon_start=1
 /product="RIKEN cdna 4632427C23 gene"
 /protein_id="AAH31933.1"
 /db_xref="GI:21594466"
 /db_xref="LocusID:78753"
 /translation="MSEILSRVTVSHRVEIWLILVAYLLQNVNSGHLPTKADPE
 AFMYSEIHKHVPSEYEVATEDGYILSVNIPRGQTRLKKEGSRPVLVLRGLLG
 DASWISNLNPNLGLFADGFDVWMSNGRNTWSRKHKTLSIDQDFWAFSYDEMA
 RPDLPVNFILQKTGOKYVGVSGQTMGFTAFSTMPDLAKHKIWMFALAPATV
 KYARSPGKFLIPDMIKVLFGQBEFYQTRFQJFYLCGMILDQICSNILLL
 GGFNTNMNRANRYVAHTPAGTSVQNLHWSOAVNSGELRAFDMGSETKQKCKNG
 PTPRYKVRDMVPTAMTGGDWLSDPDVDDVTKLLSEVTLNLIYHKNIPWAHVDFIWG
 LDAPQRYVEIHLMKQEPNLPQGTCTRVKL"

BASE COUNT 790 a 496 c 528 g 763 t

107 TAAATCGAGATGTTGGACCATGTTGGAAACCTTGTGTCAGACAGTGGATGTTGTCACACAG 166
 Db CAAATGCCGATGTTGGACCATGTCGAAATCTTGTCAAGAGTGTGGACTGTTTTCGCACAG 152
 QY AATGGAATGTTGGCTTCTGATCTGTTGGCGTATATGTTCCAGAGAAATGTGAATTCAGT 226
 Db AGTGGAGATATGGCTTCTGATCTGTTGGTAGCATATTTACTCCAAAGAAATGTGAATCTGGG 212
 QY ACATATGCCAACTPAAGCTGTGGAACCCAGAGCAATTCATGAATATAGTGAATCATCA 286
 Db ACATTTGCCCAACGAAGCTGCGGATCCAGAGCAATTCATGAATATAGTGAATCATCA 272
 QY ACATCAAGCTATCCCTGTGAGGAATATGAAGTCGCAACTGAAGATGGTATATCTCTTC 346
 Db ACAAGAGGTTATCCAGTGAAGAGATGAAGTTGCAAGAGATGGTATATCTCTTC 332
 QY TGTAAACAGGATTCCTCGAGCCCTAGTGCAACCTAAGAAAGACAGGTTCCAGCCCTGTGT 406
 Db TGTGAACAGATCCCTCGGGAACAGACACCGTTAAAGAAAGAGATCCAGGCCAGTGT 392
 QY GTTACTGCAGCATGCTAGTGTGGAGGTGTAGCACTGGATTTCCACTGCTGCCAA 466
 Db GTTACTGCACATGCTAGTGTGAGGTGTAGCACTGGATTTCCACTGCTGCCAA 452
 QY TAGCTGGGCTTCATTTCTGCGCATGCTGTTTGTGAGTGTGGTGGGAAACAGCAGAGG 526
 Db CAGCTGGGTTTATTTCTGCGCATGCTGAGTTTGTGATGTTGATGGGAAACAGCAGAGG 512
 QY AAGAGCTGTGTGCAAAACACAAAGACATCTCCATAGACCAAGATGAGTTCCTGGCTTT 586
 Db AARACCTGTGTGCGAAGCACAAGACCTCTCCATAGATCAAGATGAGTTCCTGGCTTT 572
 QY CAGTTATGATGAGATGCTAGGTTTGCACCTTCTGAGTGATGATAAACTTTATTTGAGAA 646
 Db CAGTTATGATGAGATGCTAGGTTTGCACCTTCTGAGTGATGATAAACTTTATTTGAGAA 632
 QY AACGGGCCAGAAAGATCTATTATGTCGGCTATTCAAGGGCACCACTGGCTTTAT 706
 Db AACGGGCCAGAAAGATCTATTATGTCGGCTATTCAAGGGCACCACTGGCTTTAT 692
 QY TGATTTTCCACATGCGAGCTGCTAGGTTTCAAAATCAAAATGATTTTGGCTTACCTCC 766
 Db TGATTTTCCACATGCGAGCTGCTAGGTTTCAAAATCAAAATGATTTTGGCTTACCTCC 752
 QY CATAGCCACTGTTAAGCATCAAAAGCCCGGGACCAATTTTGTGCTGCGCATAT 826
 Db TATAGCCACTGTTAAGCATCAAAAGCTCTGCTGATCCAAATTTCTGCTGCGCATAT 812
 QY GATGATCAAGGATTTTGGCAAAAGAAATTTCTGATCAGACCAAGATTTCTCAGACA 886
 Db GATGATCAAGGATTTTGGCAAGCAAGAGTTTATACCAAGACTAGATTTTTCAGGCA 872
 QY ACTTGTATTATACCTTTGCGCAGGTGATCTTTGATCAGATTTGTAGTAATATCATGTT 946
 Db GCTTTTATTATACCTTTGCGCAGATGATCTCGCAAAATCTGCGACACATCATCTT 932
 QY ACTTCTGGGTGATTCACCAACCAATATGAACATGAGCCGAGCAAGTGTATATGCTGC 1006
 Db ACTCTGGGAGGTTTAAACCAACCAATATGAACATGAGCAAGCAATGTGTATGCTGC 992
 QY CCACACTCTTCTGGAACATCTGTGCAAAATATTTCTACCTGAGCGCAGCGATGATTC 1066
 Db CCATACACTCTGGAACATCTGTGCAAGACATTTCTCCATGAGCGCGCAGCGATGATTC 1052
 QY TGGTGAATCCGGGCATTTGACTGGGGAGTGAGACCAAAATCTGGAAAAATGCAATCA 1126
 Db TGGGGAATCTGCTGCTTGAATGAGTGAAGTGAAGCAAAATCAGGAGAAATGTAATCA 1102
 QY GCCAACTCTGTAGAGTACAGATGATGAGGTCCTGAGGTCCTTACAGATGAGGACAG 1186
 Db GCCAACTCTGTAGAGTACAGATGATGAGGTCCTGAGGTCCTTACAGATGAGGACAG 1172

1187 AGCTCAGGCTGCTTCAAAATCCCAAGACAGCTGAAATGCTGCTCTGAGGTGACCAA 1246
 Db AGGTCAAGACTGCTTCAAAATCCCAAGATGATGAAATCATTTCTGAGGTGACCAA 1232
 QY CCTCATCTACCAATGAATATCTCTGAATGGGCTCAGTGGATTTTCATCTGGGTTTGA 1306
 Db CCTCATCTACCAATGAATATCTCTGAATGGGCTCAGTGGATTTTCATCTGGGTTTGA 1292
 QY TGCTCTCAGCGTATGATGAATGAATCATCTCTGATGCGAGGAGGACCAACCT 1366
 Db TGCTCTCAGCGTATGATGAATGAATCATCTCTGATGCGAGGAGGACCAACCT 1349
 QY TTCCAGGAGCGTGTGAGGCGCTATTTGGAAGCATCTGACATGACGATCTTAGACAA 1426
 Db TTCCAGGAGCGTGTGAGGCGCTATTTGGAAGCATCTGACATGACGATCTTAGACAA 1403
 QY TCCCTCAGGAGCGTGTGAGGCGCTATTTGGAAGCATCTGACATGACGATCTTAGACAA 1442
 Db TCCCTCAGGAGCGTGTGAGGCGCTATTTGGAAGCATCTGACATGACGATCTTAGACAA 1419

RESULT 4
 AL353113 43456 bp DNA linear PRI 20-SEP-2001
 LOCUS Human DNA sequence from clone Rpl1-30415 on chromosome 10. Contains
 DEFINITION the gene for a novel lipase, the gene for a novel protein similar
 to ankyrin, ESTs, STS and GSSs, complete sequence.
 ACCESSION AL353113
 VERSION AL353113.7 GI:15130723
 KEYWORDS HYG; ankyrin; lipase.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 43456)
 Bray-Allen, S.
 Direct Submission
 Submitted (23-AUG-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
 requests: clonerequest@sanger.ac.uk
 On Aug 9, 2001 this sequence version replaced gi:14669221.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e. phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given
 in the feature table with their source databases: Em: EMBL; Sw:
 SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP
 database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 10, constructed by the Sanger Centre Chromosome 10
 Mapping Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr10>
 Rpl1-30415 is from the library RPl1-11.2 constructed by the group
 of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
 VECTOR: pBAC3.6
 IMPORTANT: This sequence is not the entire insert of clone
 RPl1-30415. It may be shorter because we sequence overlapping
 sections only once, except for a short overlap.
 The true left end of clone RPl1-399019 is at 41457 in this
 sequence. The true right end of clone RPl1-186014 is at 2000 in
 this sequence.

* 28737 46598: contig of 17862 bp in length
 * 46599 gap of 100 bp
 * 46699 71912: contig of 25214 bp in length
 * 71913 gap of 100 bp
 * 72013 96612: contig of 24600 bp in length
 * 96613 gap of 100 bp
 * 96713 128331: contig of 31619 bp in length
 * 128332 128431: gap of 100 bp
 * 128432 158195: contig of 29764 bp in length.

FEATURES

source
 1..158195
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="RP11-115N17"
 /clone.lib="RPC1-11 Human Male BAC"
 1..1002
 /note="assembly_fragment"
 1103..2388
 /note="assembly_fragment"
 2489..3525
 /note="assembly_fragment"
 3626..16527
 /note="assembly_fragment"
 16628..28636
 /note="assembly_fragment"
 clone_end:T7
 vector_side:right
 28737..46598
 /note="assembly_fragment"
 46699..71912
 /note="assembly_fragment"
 72013..96612
 /note="assembly_fragment"
 96713..128331
 /note="assembly_fragment"
 128432..158195
 /note="assembly_fragment"
 clone_end:SP6
 vector_side:left
 BASE COUNT 48628 a 30917 c 30188 g 47557 t 905 others
 ORIGIN

Query Match 43.6%; Score 890.4; DB 2; Length 158195;
 Best Local Similarity 99.9%; Pred. No. 2.8e-217;
 Matches 891; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1126 AGCCAACTCCTGTAAGGTACAGAGTCAGAGATATGACGGTCCCTACAGCAATGTGGACAG 1185
 DB 19482 AGCCAACTCCTGTAAGGTACAGAGTCAGAGATATGACGGTCCCTACAGCAATGTGGACAG 19423
 QY 1186 GAGGTACAGACTGGCTTTCAATCCAGAGAGCGTGAATGCTCTCTCTGAGGTGACCA 1245
 DB 19422 GAGGTACAGACTGGCTTTCAATCCAGAGAGCGTGAATGCTCTCTCTGAGGTGACCA 19363
 QY 1246 ACCTCATCTACCATAGATATTCCTGAATGGCTCAGTGGATTTCTATCTGGGGTTGG 1305
 DB 19362 ACCTCATCTACCATAGATATTCCTGAATGGCTCAGTGGATTTCTATCTGGGGTTGG 19303
 QY 1306 ATGCTCTCACCGTATGTACATGAATCATCTCATCTGATGAGCAGGAGGAGACCAACC 1365
 DB 19302 ATGCTCTCACCGTATGTACATGAATCATCTCATCTGATGAGCAGGAGGAGACCAACC 19243
 QY 1366 TTTCACAGGAGCGGTGAGGCGGTATTGTGAGCATCTGACATGACGATCTTAGGACA 1425
 DB 19242 TTTCACAGGAGCGGTGAGGCGGTATTGTGAGCATCTGACATGACGATCTTAGGACA 19183
 QY 1426 ACCTCTGAGGATGGGGTAGGACCCATGAAGGAGCAATTCAGGAGCAGAGACCTAG 1485
 DB 19182 ACCTCTGAGGATGGGGTAGGACCCATGAAGGAGCAATTCAGGAGCAGAGACCTAG 19123
 QY 1486 TATACATTTTCAGATTCCTGCACTTGGCACTTAAATCCGACACTTACATTTT 1545

DB 19122 TATACATTTTCAGATTCCTGCACTTGGCACTTAAATCCGACACTTACATTTTACATTTT 19063
 QY 1546 TTTCCTATAAATTAAAGTACTTATTAGGTAAATAGAGGTTTGTGTATGCTATATATATTTCT 1605
 DB 19062 TTTCCTATAAATTAAAGTACTTATTAGGTAAATAGAGGTTTGTGTATGCTATATATATTTCT 19003
 QY 1606 ACCATCTTCAAGGTTAGGTTTACCTGTAGTACGAGCAATATCTAGACATTTCTCTATATC 1665
 DB 19002 ACCATCTTCAAGGTTAGGTTTACCTGTAGTACGAGCAATATCTAGACATTTCTCTATATC 18943
 QY 1666 ATTCAGTAAATCTCTTTAAACACACTATTTTCTATTAAGCCATATTTTGGAGCA 1725
 DB 18942 ATTCAGTAAATCTCTTTAAACACACTATTTTCTATTAAGCCATATTTTGGAGCA 18883
 QY 1726 CTAAAGTAAATGGCAAAATTTGGCAGATATTGAGTCTGGAGTCTGTGGATTTATTGTTG 1785
 DB 18882 CTAAAGTAAATGGCAAAATTTGGCAGATATTGAGTCTGGAGTCTGTGGATTTATTGTTG 18823
 QY 1786 ACTTTGACAAAATAAGCTAGACATTTTCACTTTGTTGCCACAGAGACATAACACTACCTC 1845
 DB 18822 ACTTTGACAAAATAAGCTAGACATTTTCACTTTGTTGCCACAGAGACATAACACTACCTC 18763
 QY 1846 AGGAAGCTGAGCTGCTTTAAGGACACACACAAATCAGTGTACAGTATGGATGAAA 1905
 DB 18762 AGGAAGCTGAGCTGCTTTAAGGACACACACAAATCAGTGTACAGTATGGATGAAA 18703
 QY 1906 TCTATGTTAAGCATTTCTCAGAAATAGGCCAAAGTTTATAGTTGTCATCTCAGGGAAGAAA 1965
 DB 18702 TCTATGTTAAGCATTTCTCAGAAATAGGCCAAAGTTTATAGTTGTCATCTCAGGGAAGAAA 18643
 QY 1966 TTTTATAGGATGTTTATGAGTTCTCATAAATGCAATCTGCAATTACATAAA 2017
 DB 18642 TTTTATAGGATGTTTATGAGTTCTCATAAATGCAATCTGCAATTACATAAA 18591

RESULT 6
 AXS21801
 LOCUS AXS21801 1116 bp DNA linear PAT 05-OCT-2002
 DEFINITION Sequence 4 from Patent WO0236731.
 ACCESSION AXS21801
 VERSION AXS21801.1 GI:23572856
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Mammalia; Euthera; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1
 AUTHORS Xiao, Y.
 TITLE Human lysosomal acid lipase
 JOURNAL Patent: WO 0236731-A 4 10-MAY-2002;
 Bayer Aktiengesellschaft (DE)
 FEATURES
 Location/Qualifiers
 1..1116
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 326 a 234 c 245 g 311 t
 ORIGIN

Query Match 36.8%; Score 752; DB 6; Length 1116;
 Best Local Similarity 97.4%; Pred. No. 7.1e-182;
 Matches 783; Conservative 0; Mismatches 0; Indels 21; Gaps 1;
 QY 273 AGTGAATCATCCAAATCAAGCTATCCCTGTGAGGATATGAGTGCACACTGAAGT 332
 DB 73 AGTGAATCATCCAAATCAAGCTATCCCTGTGAGGATATGAGTGCACACTGAAGT 132
 QY 333 GGGTATATCTTTCTGTTAAACAGGATTCCTCAGGCGCTAGTCAACCTTAAGAGACAGGT 392
 DB 133 GGGTATATCTTTCTGTTAAACAGGATTCCTCAGGCGCTAGTCAACCTTAAGAGACAGGT 192
 QY 393 TCAGGCGCTGTTGTTACTGACGATGCGCTAGTGTGAGGTTGCTAGCACTGGAATTC 452

Db 193 TCAGGCGCTGTGTGTTACTGTCAGCATGCGCTAGTTGGAGGTGCTAGCACTGGATTTC 252

Qy 453 AACCTGCCCAACATAGCTGGCTGCTTCATCTGGCAGATGCTGGTTTTCAGCTGGGATG 512

Db 253 AACCTGCCCAACATAGCTGGCTGCTTCATCTGGCAGATGCTGGTTTTCAGCTGGGATG 312

Qy 513 GGAACAGCAGGAGGAAACGCTGCTTCGAAAACACAGACACTCTCCATAGACCAAGAT 572

Db 313 GGAACAGCAGGAGGAAACGCTGCTTCGAAAACACAGACACTCTCCATAGACCAAGAT 372

Qy 573 GAGTTCTGGGCTTTCAGTTATGATGAGATGGCTAGCTTTCAGCTTTCCTGCGAGTAAAC 632

Db 373 GAGTTCTGGGCTTTCAGTTATGATGAGATGGCTAGCTTTCAGCTTTCCTGCGAGTAAAC 432

Qy 633 TTATTTTTCAGAAAACGGGCGAGGAAAGATCTATTATGTCGGCTATTTCACAGGCGACC 692

Db 433 TTATTTTTCAGAAAACGGGCGAGGAAAGATCTATTATGTCGGCTATTTCACAGGCGACC 492

Qy 693 ACCATGGGCTTTTATTCATTTTCCACCATGCTCCAGAGCTGGCTCAGAAAATCAAAATGTAT 752

Db 493 ACCATGGGCTTTTATTCATTTTCCACCATGCTCCAGAGCTGGCTCAGAAAATCAAAATGTAT 552

Qy 753 TTGCTTTTACCCATGACCTGTTTACCATGCAAAAGCCCGGACCAAAATTTTG 812

Db 553 TTGCTTTTACCCATGACCTGTTTACCATGCAAAAGCCCGGACCAAAATTTTG 612

Qy 813 TTGCTCCAGATGATGATCAAGGATTTTGGCAAAAAGAAATTTCTGTATCAGACC 872

Db 613 TTGCTCCAGATGATGATCAAGGATTTTGGCAAAAAGAAATTTCTGTATCAGACC 672

Qy 873 AGATTTCTCAGACACTGTTTATTTTACCTTTTGGCCAGGTGATTTCTGTATCAGATTGT 932

Db 673 AGATTTCTCAGACACTGTTTATTTTACCTTTTGGCCAGGTGATTTCTGTATCAGATTGT 732

Qy 933 AGTAATATCATGTTACTTCTGGTGGATTTCAACCAACCAATATGAACAT 982

Db 733 AGTAATATCATGTTACTTCTGGTGGATTTCAACCAACCAATATGAACATGATCAT 792

Qy 983 -----GAGCGAGCAAGTGTATGTCGCCACACACTTCTGTCGAAACATCTGTG 1031

Db 793 GGTGTTGTCAGAGCGCAGCAAGTGTATGTCGCCACACACTTCTGTCGAAACATCTGTG 852

Qy 1032 CAAAATATTTACACTGGAGCCAG 1055

Db 853 CAAAATATTTACACTGGAGCCAG 876

RESULT 7

AC015506/c 76653 bp DNA linear HTG 13-JUL-2000

LOCUS Homo sapiens clone RP11-21124, LOW-PASS SEQUENCE SAMPLING.

DEFINITION AC015506

ACCESSION AC015506.3 GI:9124007

VERSION HTG; HTGS_PHASE0

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 76653)

AUTHORS Birren,B., Linton,L., Nuebaum,C. and Lander,E.

TITLE Homo sapiens, clone RP11-21124

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 76653)

AUTHORS Birren,B., Linton,L., Nuebaum,C., Lander,E., Allen,N., Anderson,M., Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhgalter,B., Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArelano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M., Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Lehoczy,J., Lieu,C., Locke,K., MacDonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,

Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X., Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

Direct Submission

Submitted (16-NOV-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jul 13, 2000 this sequence version replaced gi:6468098.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RW/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L4028

Center clone name: 21_I_24

NOTE: This record contains 87 individual sequencing reads that have not been assembled into contigs. Runs of N are used to separate the reads and the order in which they appear is completely arbitrary. Low-pass sequence sampling is useful for identifying clones that may be gene-rich and allows overlap relationships among clones to be deduced. However, it should not be assumed that this clone will be sequenced to completion. In the event that the record is updated, the accession number will be preserved.

1 795: contig of 795 bp in length

796 895: gap of 100 bp

896 1656: contig of 761 bp in length

1657 1756: gap of 100 bp

1757 2558: contig of 802 bp in length

2559 2658: gap of 100 bp

2659 3434: contig of 776 bp in length

3435 3534: gap of 100 bp

3535 4301: contig of 767 bp in length

4302 4401: gap of 100 bp

4402 5199: contig of 798 bp in length

5200 5299: gap of 100 bp

5300 6085: contig of 786 bp in length

6086 6185: gap of 100 bp

6186 6972: contig of 787 bp in length

6973 7072: gap of 100 bp

7073 7844: contig of 772 bp in length

7845 7945: gap of 100 bp

7946 8731: contig of 787 bp in length

8732 8831: gap of 100 bp

8832 9625: contig of 794 bp in length

9626 9725: gap of 100 bp

9726 10519: contig of 794 bp in length

10520 10619: gap of 100 bp

10620 11417: contig of 798 bp in length

11418 11517: gap of 100 bp

11518 12296: contig of 779 bp in length

12297 12396: gap of 100 bp

12397 13180: contig of 784 bp in length

13181 13280: gap of 100 bp

13281 14065: contig of 785 bp in length

14066 14165: gap of 100 bp

14166 14949: contig of 784 bp in length

14950 15049: gap of 100 bp

15050 15822: contig of 773 bp in length

15823 15922: gap of 100 bp

15923 16719: contig of 796 bp in length

16719 16818: gap of 100 bp

16819 17621: contig of 803 bp in length

17622 17721: gap of 100 bp

17722 18517: contig of 796 bp in length

18518 18617: gap of 100 bp

	* *	51048 51147:	gap of 100 bp	
	*	51939:	contig of 792 bp in length	
	*	51940:	gap of 100 bp	
	*	52823:	contig of 784 bp in length	
	*	52824:	gap of 100 bp	
	*	52924:	contig of 791 bp in length	
	*	53715:	gap of 100 bp	
	*	53815:	contig of 790 bp in length	
	*	54605:	gap of 100 bp	
	*	54705:	contig of 787 bp in length	
	*	55491:	gap of 100 bp	
	*	55591:	contig of 794 bp in length	
	*	55592:	gap of 100 bp	
	*	56386:	contig of 779 bp in length	
	*	56486:	gap of 100 bp	
	*	57264:	contig of 780 bp in length	
	*	57365:	gap of 100 bp	
	*	58145:	contig of 780 bp in length	
	*	58245:	gap of 100 bp	
	*	59001:	contig of 757 bp in length	
	*	59101:	gap of 100 bp	
	*	59102:	contig of 774 bp in length	
	*	59876:	gap of 100 bp	
	*	59976:	contig of 778 bp in length	
	*	60754:	gap of 100 bp	
	*	60854:	contig of 783 bp in length	
	*	61637:	gap of 100 bp	
	*	61737:	contig of 772 bp in length	
	*	62509:	gap of 100 bp	
	*	62609:	contig of 774 bp in length	
	*	63383:	gap of 100 bp	
	*	63483:	contig of 778 bp in length	
	Query Match Score 630.2; DB 2; Length 76653;			
	Best Local Similarity 99.4%; Pred.No. 1.7e-150;			
	Matches 632; Conservative 0; Mismatches 4; Indels 0; Gaps 0;			
QY	1126	AGCCAACTCCTGTAAGTACAGAGTCAGAGTGATGACGGTCTTNACAGCAATTGTGGACAG	30665	
DB	30724	AGCCAACTCCTGTAAGTACAGAGTCAGAGTGATGACGGTCTTNACAGCAATTGTGGACAG	30665	
QY	1186	GAGGTCAAGACTGCCTTTCAAATCCAGAGAGCGTGAATAAATGCTGCTCTCAGGTGACCA	1245	
DB	30664	GAGGTCAAGACTGCCTTTCAAATCCAGAGAGCGTGAATAAATGCTGCTCTCAGGTGACCA	30605	
QY	1246	ACCTCATCTACCATAAGAATAATTCCTGAATGGGCTCACGTGATTTCACTCGGGGTTGG	1305	
DB	30604	ACCTCATCTACCATAAGAATAATTCCTGAATGGGCTCACGTGATTTCACTCGGGGTTGG	30545	
QY	1306	ATGCTCCTCACCGTAGTACAAATCAATCAATCATCTGATGACAGAGAGAGACCAACC	1365	
DB	30544	ATGCTCCTCACCGTAGTACAAATCAATCAATCATCTGATGACAGAGAGAGACCAACC	30485	
QY	1366	TITCCCAGGAGCGGTGTAGGCCGTATTGTGAAGCATCTGACATGAGCATCTTAGGACA	1425	
DB	30484	TITCCCAGGAGCGGTGTAGGCCGTATTGTGAAGCATCTGACATGAGCATCTTAGGACA	30425	
QY	1426	ACCTCCTGAGGAGTGGGGCTAGSACCATGAAAGGAGAAATTCGAGAGAGAGACCTTAG	1485	
DB	30424	ACCTCCTGAGGAGTGGGGCTAGSACCATGAAAGGAGAAATTCGAGAGAGAGACCTTAG	30365	
QY	1486	TATACATTTTTTCAGATTCCCTGCACCTTGGCAGTAAATCGCACATTTACATTTT	1545	
DB	30364	TATACATTTTTTCAGATTCCCTGCACCTTGGCAGTAAATCGCACATTTACATTTT	30305	
QY	1546	TTTCTGTAATTTAAAGPACTTATTAGGTAANAATAGAGGTTTTGTATGCTATTATATTCT	1605	
DB	30304	TTTCTGTAATTTAAAGPACTTATTAGGTAANAATAGAGGTTTTGTATGCTATTATATTCT	30245	
QY	1606	ACCATCTTGAAGGTAGGTTTTTACCTGATAGCCAGAAAATCTAGACATTTCTTATATC	1665	
DB	30244	ACCATCTTGAAGGTAGGTTTTTACCTGATAGCCAGAAAATCTAGACATTTCTTATATC	30185	
QY	1666	ATTGAGTAAATCTCTTTAAACACCTATTGTTTTTTCTTATAAGCCATTTTTTTGGAGCA	1725	

Db 30184 ATTCAAGTAAATCTCTTTAAACACCTATTGTTTTTCTATAGCAATATTTTGGAGCA 30125

Qy 1726 CTAAAGTAAATGGCAAAATGGGACAGATATTGAGG 1761

Db 30124 CTAAAGTAAATGGCAAAATGGGACAGATATTGAGG 30089

RESULT 8

AK091558

LOCUS

DEFINITION Homo sapiens cDNA FLJ34239 fis, clone FCBBF3027755, highly similar to LYSOSOMAL ACID LIPASE/CHOLESTERYL ESTER HYDROLASE PRECURSOR (EC 3.1.1.13).

ACCESSION AK091558.1 GI:21749961

VERSION

KEYWORDS oligo capping; fis (full insert sequence).

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE

AUTHORS Ishibashi, T., Kanehori, K., Yosida, M., Watanabe, S., Ishida, S., Ono, Y., Houta, T., Hiraoka, S., Murakawa, K., Takiguchi, S., Kusano, J., Watanabe, M., Fujimori, K., Tanai, H., Ishida, M., Yamashita, H., Chiba, Y., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Ota, T., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuho, Y., Nagai, K. and Isogai, T.

TITLE NEDO human cDNA sequencing project

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 2982)

AUTHORS Isogai, T. and Yamamoto, J.

TITLE Direct Submission

JOURNAL Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)

COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5' & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES

source

1. .2982

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="FCBBF3027755"

/tissue_type="brain"

/clone_lib="FCBBF3"

/dev_stage="fetal"

/notes="cloning vector: pME18SFL3"

BASE COUNT 847 a 596 c 615 g 924 t

ORIGIN

Query Match 25.8%; Score 528.2; DB 9; Length 2982;

Best Local Similarity 65.4%; Pred. No. 2.2e-124;

Matches 792; Conservative 0; Mismatches 413; Indels 6; Gaps 1;

Qy 162 CACAGATGAATGGCTTCTGATTCCTGGTGGCGTATATGTTCCAGAGAAATGTGAAT 221

Db 672 CACAGATGAATGGCTTCTGAGGTTGGTGGTCTGTTGGTCTCTGGACCTGGAT 731

Qy 222 TCAG-----TACATATGCCAATAAAGCTGTGGACCCAGAGCATTCATGAATATTAGT 275

Db 732 TCTGAGGGTCTGGAGGGAACCTGACAGCTGTGGATCTCTCAACAAACATGAATGTAGT 791

Qy 276 GAAATCATCAACATCAAGCTATCCCTGTGGAGGAATATGAGTCCCACTGAAGATGGG 335

Db 792 GAAATCATCTTTACTGGGATTTCCCTAGTGAGGAATACTAGTTGAGACAGAGATGGA 851

Qy 336 TATATCCTTTCTGTGTTAAACAGGATTTCTCGAGGCTAGTGCACCTTAAGAGACAGGTTCC 395

Db 852 TATATCTGTGCTCTTAACCGAATTTCTCTATGGGAGGAAGAACATTTCTGACAAAGTCCC 911

Qy 396 AGGCTGTGCTGTGTTACTGACAGCATGGCTAGTTGGAGGTGCTAGCAACTGGATTTCCAAAC 455

Db 912 AAACCAAGTTCTCTCTGCAACATGGCTGCTGCGCAGATTTAGTAACCTGGGTCAACAAC 971

Qy 456 CTGCCCAACAAATAGCTGGCTTCATCTGCGCAGATGCTGGTTTTCACGTGTGGATGGGG 515

Db 972 CTTGCCAGCAGCAGCTGGGCTTCATTTCTGCTGATGCTGGTTTTCACGTGTGGATGGGC 1031

Qy 516 AACAGCAGGGGAAACGGCTGCTCGAAACACACAGACACTCTCCATAGACCAAGATGAG 575

Db 1032 AACAGCAGAGGAATACCTGGTCTCGGAACATCAGACACTCTCAGTTCTCAGGATGAA 1091

Qy 576 TTCTGGGCTTTCAAGTTATGATGAGTGGCTAGGTTTGGACCTTCTCGCAGTGATAAATTT 635

Db 1092 TTCTGGGCTTTCAAGTTATGATGAGTGGCAAAATATGACCTACCACTTCCATTAACTTC 1151

Qy 636 ATTTTGACAGAAACGGGCCAGGAAAGATCTATTATGTGGCTATTTCACAGGGCACACC 695

Db 1152 ATTCTGAATAAACTGGCCAGAAACAGTGTATATGTGGTCTATCTCAGGCACCACT 1211

Qy 696 ATGGGCTTTATTCATTTTCCACCATGCGCAGAGCTGGCTCAGAAAAATCAAAATGTTT 755

Db 1212 ATAGTTTTTATAGCAATTTTACAGATCCCTGAGCTGGCTAAAAGGATTAATAATGTTTT 1271

Qy 756 GCITTAGCACCCATAGCACTGTTAGCATGCAAAAGCCCGGACCAAAATTTTGTG 815

Db 1272 GCCCTGGGCTCTGTGGCTTCCGTCGGCTTCTGTACTAGCCCTATGCGCCAAATTAGGACGA 1331

Qy 816 CTGCCAGATATGATGATCAAGGATTTGTTGGCAAAAGAAATTTCTGTATCAGACCGA 875

Db 1332 TTACCAATCATCTCATTAAGGACTTATTTGGAGACAAAGAAATTTCTTCCCGAGGTGG 1391

Qy 876 TTCTCAGACAACTGTTATTTACCTTTGTTGGCCAGGTGATTTCTGATCAGATTTAGT 935

Db 1392 TTTTGAAGTGGGTGGGTACCCACGTTTGCACTCATGCTAAGGAGCTCTGTGGA 1451

Qy 936 AATATCATGTTACTTCTGGGTGATTCACACCAACAATATGAACATGAGCCGCAAGT 995

Db 1452 AATCTCTGTTTCTCTGTGTGGATTTAATGAGAGAAATTAATATGCTAGAGTGGAT 1511

Qy 996 GTATATGCTGCCACACTCTGTTGCGAAACATCTGTGCAAAATATTTCTACCTGAGCCAG 1055

Db 1512 GTATATCAACACATTTCTCTGCTGGAACTTCTGTGCAAAACATGTTACACTGGAGCCAG 1571

Qy 1056 GCAGTGAATTTCTGTGAATCTCGGGCATTTGACTGGGGAGTGAGACCAAAATCTGGAA 1115

Db 1572 GCTGTAAATTCGAAAGTTTCAAGCTTTGATGCGGAGCAGTGCAGGAATCATTTT 1631

Qy 1116 AAATGCAATCAGCAACTCTCTGTAAGTACAGAGTACAGATATGACGGTCCCTACAGCA 1175

Db 1632 CATTACCAACAGAGTTATCTCTCCACATACAAATGTGAAGACATGCTGTGCGCAGTCGA 1691

Qy 1176 ATGTGACAGAGTGCAGGCTGCTTCAATCCAGAGAGCTGGAATGCTGCTCT 1235

Db 1692 GTCTGGAGCGGGGTGACGACTGGCTTGCAGATGTTCTACGACGTCAATATCTTACTGACT 1751

Qy 1236 GAGGTGACCAACCTCATCTACCATAGAAATATTCTCTGMAATGGGCTCACGTGGATTTTATC 1295

Db 1752 CAGATCACCACCTGGTGTTCATGAGAGCATTCGGAATGGGAGCATCTTGACTTTTCAAT 1811

Qy 1296 TGGGTTTGGATGCTCTCTACCGTATGTACAAATGAATCATCTCATGTGAGCAGGAG 1355

Db 1812 TGGGGCTTGGATGCTCTCTACCGTATGTACAAATGAATCATCTCATGTGAGCAGGAG 1871

Qy 1356 GAGACCAACCT 1366

Db 1872 CAGTCAAGCT 1882

```

RESULT 9
AX411234
LOCUS AX411234 2481 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 3881 from Patent WO0229103.
ACCESSION AX411234
VERSION AX411234.1 GI:21443939
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J.G.
Gene expression profiles in liver cancer
Patent: WO 0229103-A 3881 11-APR-2002;
GENE LOGIC INC (US)
FEATURES
source
1..2481
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. Z31690"
BASE COUNT 692 a 477 c 513 g 799 t
ORIGIN
Query Match 25.8%; Score 527.8; DB 6; Length 2481;
Best Local Similarity 66.6%; Pred. No. 2.7e-124;
Matches 751; Conservative 0; Mismatches 376; Indels 0; Gaps 0;
QY 240 AAGCTGTGGACCCAGAGAGCTTATGAATATTAGTGAATATCCACATCAAGGCTAT 299
DB 112 ACAGCTGTGGATCCTGAAACAAACATGAATGTGAGTGAATATCTCTTACTGGGATTC 171
QY 300 CCTGTGAGGAATATGAAGTGGCACTCAAGATGGGTATATCTTCTGTAAACAGATT 359
DB 172 CCTAGTAGGAATACCTAGTTAGACAGAGATGGAATATCTGTGCTTAAACCGAATT 231
QY 360 CTTGAGGCTTGTGCAACCTAAGAGACAGGTTCCAGGCTGTGGTGTACTGACGAT 419
DB 232 CTTGAGGAGGAGAGAACCATCTTGACAAAGGTCCCAACAGTGTCTTCTCTGCAACAT 291
QY 420 GGCCTAGTTGGAGGCTGACGATGATGATTTCCACCTGCCCAACATAGCTGGGCTTC 479
DB 292 GGCCTGCTGGCAGATCTTGAATCTGAGTGGGTCAACCTTGCACACAGAGCTGGGCTTC 351
QY 480 ATTCTGGCAGATGCTGGTGTGATGGGAGACAGAGGAGAAAGCGCTGGTCT 539
DB 352 ATCTTGTGATGCTGGTGTGATGGGAGACAGAGGAGAAATACCTGGTCT 411
QY 540 CGAAACACAGACACTCTCCATGACCAAGATGATGTTCTGGGCTTTCAGTTATGATGAG 599
DB 412 CGAAACACATAAGACACTCTCAGTTTCTCAGGATGAATTTCTGGGCTTTCAGTTATGATGAG 471
QY 600 ATGGCTAGTTGACCTTCTGCTGAGTGAATTTATTTTGCAGAAAGCGGCGCAAGAA 659
DB 472 ATGGCAAAATATGACTACAGCTTCCATTAATCTTCAATTAATAATGCGCAAGAA 531
QY 660 AAGATCTATTATGTCGGCTATTTCAGAGGCGCCACCAATGGGCTTTATTCGATTTTCCACC 719
DB 532 CAAGTGTATTATGTTGGTCAATCTCAAGGCGCCACTATAGTGTATTATAGCATTTTCACAG 591
QY 720 ATGCCAGAGCTGGCTCAGAAATCAAAATGATTTTGTGTTAGCACCATAGCCACTGTT 779
DB 592 ATCCCTGAGCTGGCTTAAAGGATTAATTAATGTTTGTGCTGGTCTGCTGGCTTCCGTC 651
QY 780 AAGCATGCAAAAGCGCGGACCAAAATTTTGTGCTGCCAGATATGATGATCAAGGGA 839
DB 652 GCCTTCTGTACTAGCCCTATGGCAAAATAGGACGATTAACGATCATCTCATTAAGGAC 711
QY 840 TGTGTTGCCAAAAAGAAATTTCTGATCAGACAGATTTCTCAGACAACTGTTATTATAC 899

```

```

DB 712 TTATTGGAGACAAAGAAATTTCTCCCGAGAGTGGTTTTTCAAGTGGCTGGTACCAC 771
QY 900 CTTTGGGCCAGGTGATTTCTTGATCAGATTTGTAGTAATATCATGTTACTTCTGGGTGGA 959
DB 772 GTTTGACATCATGTCTATCTGAAGGAGCTCTGTGAAATCTCTGTTTCTTCTGTGTGGA 831
QY 960 TTCAACACCAACAAATATGAACATGAGCCGAGCAAGTGTATATGTCGCCACACTCTTCT 1019
DB 832 TTTAATGAGAGAAATTTAAATATGTCTAGAGTGGATATATACACACATCTCTCTCT 891
QY 1020 GGAACATCTGTGCAAAATATTCTACACTGGAGCCAGGAGTGAATTTCTGGTGAATCCGG 1079
DB 892 GGAACCTTCTGTGCAAAACATGTTACACTGGAGCCAGGCTGTTAAATTCGAAAGTTTCAA 951
QY 1080 GCATTTGACTGGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCAACTCTCTGTA 1139
DB 952 GCCTTTGACTGGGAGAGCGTCCAGAAATATTTTCAATACACAGAGTATCTCTCC 1011
QY 1140 AGGTACAGAGTCAGAGATATGACGGTCCCTACAGCAATGTGGACAGGAGTCCAGACTGG 1199
DB 1012 ACATACATGTGAAGGACATGCTTGTGCGACTGCACTGTGGAGCGGGGTCCAGACTGG 1071
QY 1200 CTTTCAATCCAGAGAGCGTGAATGCTCTCTGAGGTGACCAACTCATCTACCAT 1259
DB 1072 CTTGCAAGTGTCTACGAGTCAATATCTTACTGACTCAGATCACCACCTTGGTGTCCAT 1131
QY 1260 AAGATATTCTGAAATGGGCTCACGCTGATTTTCATCTGGGTTTGGATGCTCTCACCGT 1319
DB 1132 GAGAGCATTTCCGAAATGGGAGCATCTTGACTTTCATTTGGGCTGATGCCCTTGGAGG 1191
QY 1320 ATGTACATGAATCATCTCATCTGATGACGAGGAGGAGACCAACT 1366
DB 1192 CTTTAAATAAATATTAAATCTTAATCTAATGAGGAATATCAGTGAAGCT 1238

RESULT 10
HSLYACLY
LOCUS HSLYACLY 2481 bp mRNA linear PRI 17-FEB-1997
DEFINITION H.sapiens (HepG2) LAL mRNA for lysosomal acid lipase.
ACCESSION Z31690
VERSION Z31690.1 GI:506430
KEYWORDS LAL; lipase; lysosomal acid lipase.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. Du, H. and Gregory, G.A.
Structural Conservation of Putative Functional Motifs between Mouse
and Human Lysosomal Acid Lipase
Unpublished
REFERENCE
2. (bases 1 to 2481)
Du, H.
Direct Submission
Submitted (05-APR-1994) Hong Du, Division of Human Genetics,
Children's Hospital Medical Center, 3333 Burnet Street, Cincinnati,
OH, 45229-3032, USA
FEATURES
source
1..2481
/organism="Homo sapiens"
/mol_type="mRNA"
/strain="HepG2"
/db_xref="taxon:9606"
/clone="clone hLAL"
/cell_type="hepatoma"
/tissue_type="liver"
/clone_lib="human liver HepG2 cDNA library/lamba ZAP"
1..2481
/gene="LAL"
34..1233
/gene="LAL"
/citation=[1]
/codon_start=1

```

```

/product="lysosomal acid lipase"
/protein_id="CAA83495.1"
/db_xref="GI:506431"
/translation="MKQRFGLVPLVCLVLTLSHSGRGLTAVDPETNMVSEIISYW
GPESEYLVETEDGYILCLNRIPIHGRKNHSDKPKPVVFLQHLGLADSSNWNVLANS
SLGFIILADAGPDMGNSRGNTWSRKHKTLVSQDFWAFSDYDEMAKVDLPASINFL
NKTGQOVYVGVHSGQTTIGFIAPSOIPELAKRIKMFALGPVAVAFCTSPMAKIGR
LPDLHLKDLFGKEFLPQSAFLKWLGVHCTHVLKELCGNLCELLCGFERNLNMSR
VDYTHSPAGTSVQNLHWSQAVKFKQFADMGSSAKNYHFNQSPPTNYNVDML
VPTAVWGGHMDLADYDVNILLTQITLVFPHESIPWEHEDFINGLDAPRLNKKII
NLMRYQ"
34..96
sig_peptide
/gene="LAL"
/citation=[1]
97..1230
mat_peptide
/product="lysosomal acid lipase"
/citation=[1]
79
variation
/gene="LAL"
/citations=[1]
/replace="c"
100
variation
/gene="LAL"
/citations=[1]
/replace="g"
2459..2464
polyA_signal
/gene="LAL"
2480
polyA_site
/gene="LAL"
692 a 477 C 513 G 799 T
BASE COUNT
ORIGIN
Query Match 25.8%; Score 527.8; DB 9; Length 2481;
Best Local Similarity 66.6%; Pred. No. 2.7e-124;
Matches 751; Conservative 0; Mismatches 376; Indels 0; Gaps 0;
QY 240 AAAGCTGTGAGCCAGAGCAATTCATGAATATTAGTGAATATCAACATCAAGGCTAT 299
DB 112 ACAGCTGTGGATCTCGAACAACATGAATGTGAGTGAATATCTCTACTGGGATTC 171
QY 300 CCCTGTGAGGAATATGAAGTCGCACTGAAGTGGTATATCTCTTCTGTAAACAGATT 359
DB 172 CCTAGTGAAGTAATACCTAGTTGAGACAGAGATGATATCTGTGCGCTTAACCGAAT 231
QY 360 CCTCGAGGCTTAGTGCACTTAAGACAGACAGGTTCCAGGCTGTGGTGTTACTGCAAGCAT 419
DB 232 CCTCATGGGAGGAAGAACCATTCGACAAAGTCCCAACAGTTGTCTTCTCTGCAACAT 291
QY 420 GGCTAGTGTGAGGTGTAGCACTGAGATTTCCAACTGCGCCACATAGCTGGGCTTC 479
DB 292 GGCTGTGCGGAGATTCAGTAATCTGGGTCAAAACCTTGCAACAGACGCTGGGCTTC 351
QY 480 ATTCTGGCAGATGCTGTTTTCAGCTGTGGATGGGGAACAGCAGGGGAACGCTGTCT 539
DB 352 ATTCTGTGATGCTGTTTTCAGCTGTGGATGGGGAACAGCAGGGAATACCTGTCT 411
QY 540 CGAAACACAGACACTCTCATAGACCAAGATGATGTTCTGGGCTTTCAGTTATGATGAG 599
DB 412 CGAAACACAGACACTCTCATAGACCAAGATGATGTTCTGGGCTTTCAGTTATGATGAG 471
QY 600 ATGGCTAGGTTTTCAGCTCTTCAGTGAATAACTTTATTTTGCAGAAACGGGCGAGAA 659
DB 472 ATGGCAAAATATGACCTACAGCTCTCAATTAACCTTCAATCTGATATAAATCTGGCCAGAA 531
QY 660 AAGATCTATTATGTCGGGTATTCACAGGCAACCACTAGGCTTTATGCAATTTCCACC 719
DB 532 CAAGTGTATTATGTCGGGTATTCACAGGCAACCACTAGGCTTTATGCAATTTTCACAG 591
QY 720 ATGCCAGAGCTGGCTCAAAATCAAAATGATTTTGTGTTTATGACCCATAGCCACGTGT 779
DB 592 ATCCCTGAGCTGGCTAAAGAGGATTAATGATTTTTTGGCTGGGTCTGTGGCTTCCGTC 651

```

QY 780 AAGCATCAAAAAGCCCGGAGCAAAATTTTCTGCTGCCAGATATGATCAAGGGA 839

DB 652 GCCTTCTGTACTAGCCCTTGGCCAAATTAAGAGAGATTACAGATCATCTCATTAAGAC 711

QY 840 TTGTTTGGCAAAAAGAAATTTCTGTATCAGACCAAGATTTCTCAGACAACTTGTATTAC 899

DB 712 TTAATTTGAGACAAAGAAATTTCTCCCAAGAGTGGCTTTTGAAGTGGGTGGGTACCCAC 771

QY 900 CTTTGTGGCCAGGTGATTTCTTGATCAGATTTGTAGTAATATCATGTTACTTCTGGGTGA 959

DB 772 GTTTGCACTCATGTCACTAGAGAGCTCTGTGGAAATCTCTGTTTTCTTCTGTGTGA 831

QY 960 TTCAACACCAACATATGAACATGAGCCGAGCAAGTGTATATCTGCCCACTCTTGTCT 1019

DB 832 TTTAATGAGAGAAATTTAAATATGTCTAGAGTGGATATATATACAAACACATTTCTCTGCT 891

QY 1020 GGAACATCTGTGCAAAATATTTCTACCTGAGCCAGCAGTGAATTTCTGTGTAACCTCCG 1079

DB 892 GGAACCTTCTGTGCAAAACATGTTACCTGGAGCCAGCTGTAAATTTCCAAAAGTTTCAA 951

QY 1080 GCATTGTGACTGGGGAGTGAGACCAAAATCTGAAAAATGCAATCAGCCAACTCTCTGTA 1139

DB 952 CCTTTGACTGGGAGCAGTGCAGAAATTTATTTTCAACACAGAGTTATCTCTCC 1011

QY 1140 AGGTACAGAGTCAGATATGACGGTCCCTACAGCAATGTGGACAGAGGTGAGGACTGG 1199

DB 1012 ACATACAAATGTGAAGGACATGCTTTGTGCCAGCTSCAGTCTGGAGCGGGGTACGACTGG 1071

QY 1200 CTTTCAAAATCCAGAGAGACGTGAAATGCTGCTCTCAGGTGACCAACCTCATCTACCAT 1259

DB 1072 CTTGCAATGTCTACGAGCTCAATATCTTACTGACTCAGATCACCACCTTGGTGTCCAT 1131

QY 1260 AAGAAATTTCTGAATGGCTCAGCTGATTTCACTGGGGTTTGGATGCTCTCACCCT 1319

DB 1132 CAGAGCATTCGGGAATGGGAGCATCTTCACTTCAATTTGGGGCCCTGGATGCTTGGAGG 1191

QY 1320 ATGTACATGAATCAATCCATCTCATGACGAGGAGGAGACCACT 1366

DB 1192 CTTTATAATAAATTTAATTAATGAGAAATATCAGTGAAAGCT 1238

RESULT 11

HSU08464

LOCUS Human lysosomal acid lipase mRNA, complete cds.

DEFINITION

ACCESSION U08464

VERSION U08464.1

KEYWORDS GI:505052

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2481)

AUTHORS Du,H. and Gregory,G.A.

TITLE Structural Conservation of Putative Functional Motifs between Mouse and Human Lysosomal Acid Lipase

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 2481)

AUTHORS Du,H.

TITLE Direct Submission

JOURNAL Submitted (08-APR-1994) Hong Du, Division of Human Genetics, Children's Hospital Medical Center, 3333 Bernet Street, Cincinnati, OH 45229-3039, USA

FEATURES

Location/Qualifiers

1..2481

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="clone hLAL"

/cell_line="HepG2"

/cell_type="hepatoma"

/tissue_type="liver"


```

/clone_lib="HepG2 cdna library/Lambda ZAP"
1. .33_
34. .1233
/codon_start=1
/product="lysosomal acid lipase"
/protein_id="AAB60328.1"
/db_xref="GI:505053"
/translating="MKMRFLGLVCLVLTWLTSHSGSRGKLTAVDPETNMVSRILSYW
GFPSEVLVEDDGYILCLNRIPEGRKNHSDKPKVPVLOHGLLADSSNWTNLANS
SLKTEILADAGDVDMGMSRGNWRSKHTLSVQDEFWAFSDYEMAKYDLPASINFIL
NKTGEQVIVHSGGHTTIGTIAFQIPELAKRIKFPFALGPVAVAFCTSPNAKUGR
LPDLHILKDLFDEKFLPQSAFLKWLGHVTHVLKELCNLCLFLGCFERNLNMNR
VDVTTSPAGTSQFQMLHMSQAVKFKQFQAFDMSSAKNYFHYNSYPTVYNKMDL
VPTAVSGGSHDLADVDVYNILLTQITNLVFHESIPWEHLDFIWLGLDAPWRLYNKII
NLMEKYQ"
34. .95
sig_peptide
mat_peptide
variation
79
/product="lysosomal acid lipase"
/replacement="c"
100
/replacement="g"
1234. .2480
polya_signal
2459. .2464
polya_site
2480
BASE COUNT 692 a 477 c 513 g 799 t
ORIGIN

Query Match 25.8%; Score 527.8; DB 9; Length 2481;
Best Local Similarity 66.6%; Pred. No. 2.7e-124;
Matches 751; Conservative 0; Mismatches 376; Indels 0; Gaps 0;

QY 240 AAAGCTGTGGACCCAGAGCAATCATGAATATTAGTGAATCATCCAACTCAAGGCTAT 299
DB 112 ACAGCTGTGGATCCTGAAACCAACATGATGAGTGAATATCTCTTACTGGGATTC 171
QY 300 CCCTGTGAGGATATGAGTGCACACTGAAGTGGTATATCTCTTGTATACAGGATT 359
DB 172 CCTAGTGGGAATACCTAGTGTGAGACAGAGATGGATATATCTGTGCCCTTAACCGAAT 231
QY 360 CCTGAGGACCTAGTGCAACCTAAGAAGACAGGTTCCAGGCTGTGGTGTACTGACGAT 419
DB 232 CCTCATGGGAGGAGAACCACTTCTGACAAAGGTCCCAACCCAGTGTCTTCTTCTGCAACAT 291
QY 420 GGCTAGTGTGAGGCTAGCACTGATTTTCCAACTGCTGCCCAACAATAGCTGGGCTTC 479
DB 292 GGCTGTCTGGCAGATCTTAGTAACTGGGTGCTCAAAACCTTGCACACAGCAGCTGGGCTTC 351
QY 480 ATTCTGCAGATGCTGGTGTGAGTGGGATGGGAAACAGCAGGGAAGCGCTGGTCT 539
DB 352 ATTCTGTGATGCTGGTGTGAGTGGGATGGGAAACAGCAGGGAATACCTGGTCT 411
QY 540 CGAAACACAGACACTCTCCATAGACCAAGATGAGTGTGGGCTTTCAATATGATGAG 599
DB 412 CGAAACACAGACACTCTCAGTGTCTCAGGATGAATCTGGGCTTTCAATATGATGAG 471
QY 600 ATGGCTAGGTTGACCTTCTGAGTGAATACCTTTATTTTGAGAAACGGGCGAGAA 659
DB 472 ATGGCAAAATGAGCTTACCTACCTGCTTCCATTAACCTTCTGGAATAAATCGGCAAGAA 531
QY 660 AAGATCTATTATGTCGGCTATTTCAGGCGCACCACTGGCTTTATTTGCAATTTCCACC 719
DB 532 CAAGTGATATTATGTCGGCTATTTCAGGCGCACCACTATAGGTTTATAGCAATTTTCACAG 591
QY 720 ATGCCAGAGCTGGCTCAGAAATCAAAATGATTTTGTCTTTAGCAACCATAGCCAGCTTT 779
DB 592 ATCCCTGAGCTGGCTTAAAGGATTAATGATTTTGTCCCTGGCTCTGGCTTCGGCTC 651
QY 780 AAGCATGCAAAAGAGCCCGGACCAAAATTTTGTCTGCCAGATATGATGATCAAGGCA 839
DB 652 GCCTTCTGACTAGCCCTATGCCAAATAGGACGATTAACGATCATCTCATTAAAGAC 711
QY 840 TTGTTTGGCAAAAAGAAATTTCTGTATCAGACCGAGATTTCTCAGAACACTTGTATTATTC 899

```

5'UTR
CDS

Db 712 TTATTGGAGACAAAGAAATTTCTCCCGAGAGTGGTTTTTGAAGTGGCTGGTACCCAC 771

QY 900 CTTTGTGGCCAGGTGATTTCTTGATCAGATTTGTAGTAATATATCATGTTACTTCTGGTGGGA 959

Db 772 GTTTCGACTCATGTCTATCTGAAGGAGCTCTGTGAAATCTCTGTGTTTCTTCTGTGTGGA 831

QY 960 TTCAACACCAACAATATGAACATGATGAGCCGAGCAAGTGTATATGCTGCCACACTCTTCT 1019

Db 832 TTTATGAGAGAAATTTTAAATATGTTCTAGTGGATGTATATACAAACATTTCTCTCT 891

QY 1020 GGAACATCTGTGCAAAATATTCTACACTGGAGCCAGGAGTGAATTTCTGGTGAATCTCGG 1079

Db 892 GGAATCTTCTGTGCAAAACATGTTTACACTGGAGCCAGGCTGTAAATTTCCAAAGTTTCAA 951

QY 1080 GCATTGTGCTGGGAGTGAAGCAACCAAAATCTGGAATAATGCAATCAGCAACTCTCTCTA 1139

Db 952 GCCTTTGAGCTGGGAGCAGTGGCCAAAGATTTATTTTCAATCAACGAGTATCTCTCCC 1011

QY 1140 AGGTACAGAGTCAGAGATATGACGGTCCCTACAGCAATGTGGACAGGAGTCAAGACTGG 1199

Db 1012 ACATCAATGTGAAGACATGCTTGTGCGGACTGCAGTCTGGAGCGGGGTCAAGACTGG 1071

QY 1200 CTTCAATTCAGAGAGCTGAAATGCTGCTCTGTAGGTGACCAACTCATCTACCAT 1259

Db 1072 CTTGCAGATGCTCTACGAGCTCAATATCTTACTGACTCAGATCACCACCTTGGTGTTCAT 1131

QY 1260 AAGATATTTCTGAAATGGGCTCAGCTGAGTATTTCACTGGGGTTTGATGCTCCTCACCGT 1319

Db 1132 GAGAGCATTTCCGAGTGGAGCATCTTGCATTTATTTGGGCCCTGGATGCCCTTGGAGG 1191

QY 1320 ATGTCAATGAATCATCATCTGATGAGCAGGAGGAGAGCAACCT 1366

Db 1192 CTTTATAATAAATTTAATCTAATGAGAAATATCAGTGAAGCT 1238

RESULT 12

HSLAL

LOCUS HSLAL 2626 bp mRNA linear PRI 25-FEB-1994

DEFINITION H.sapiens mRNA for lysosomal acid lipase.

ACCESSION X76488

VERSION X76488.1 GI:434305

KEYWORDS lysosomal acid lipase.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1

Ameis,D., Merkel,M., Eckerskorn,C. and Greten,H.
Purification, characterization and molecular cloning of human
hepatic lysosomal acid lipase
Eur. J. Biochem. 219 (3), 905-914 (1994)

JOURNAL MEDLINE 94155897

PUBMED 8112342

REFERENCE 2 (bases 1 to 2626)

Ameis,D.

Direct Submission

Submitted (29-NOV-1993) D. Ameis, Medical Department, University
Hospital Eppendorf, Martinistrasse 52, 20246 Hamburg, FRG

LOCATION/Qualifiers

source

1. .2626
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosomes="10"
/tissue_type="liver"
1. .2626
/gene="LAL/LIPA"
146. .1345
/gene="LAL/LIPA"
/EC_number="3.1.1.13"
/note="sterol esterase"
/codon_start=1

gene

CDS

Db 1079 CTTGCAGATGTCTACGACGCTCAATATCTTACTGACTCAGATCACCACCTTGTTGTTCCAT 1138

Qy 1260 AAGAATATTCTCGAATGGGCTCAGCTGGAGTTTCATCTGGGTTTGGATCCTCCTCACCGT 1319

Db 1139 GAGAGCATTCGGATGGAGGATCTTGATCTTCATTTGGGGCTGGATGCCCTTTGGAGG 1198

Qy 1320 ATGTACAATGAATCATCATCTGATGCAGCAGGAGGAGACCAACT 1366

Db 1199 CTTTATAATAAATTATTAACTTAATCAGGAAATATCAGTGAAGCT 1245

RESULT 14

HMLIPCHL 2493 bp mRNA linear PRI 07-JAN-1995

LOCUS Human lysosomal acid lipase/cholesterol esterase mRNA, complete cds.

DEFINITION

ACCESSION M74775

VERSION 1 GI:187151

KEYWORDS lysosomal acid lipase/cholesterol esterase.

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 2493)

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

TITLE Anderson, R.A. and Sando G.N.

Cloning and expression of cDNA encoding human lysosomal acid lipase/cholesterol ester hydrolase. Similarities to gastric and lingual lipases

JOURNAL J. Biol. Chem. 266 (33), 22479-22484 (1991)

MEDLINE 92042192

PUBMED 1718995

COMMENT Original source text: Homo sapiens fibroblast cDNA to mRNA.

FEATURES

Location/Qualifiers

1..2493

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/map="10"

/tissue_type="fibroblast"

1..2493

/gene="LIPA"

<1..2493

/gene="LIPA"

/note="G00-120-153"

41..1240

/gene="LIPA"

/EC_number="3.1.1.13"

/note="alternate initiator ATG may be at position 47"

/codon_start=1

/product="lysosomal acid lipase/cholesterol esterase"

/protein_id="AAA59519.1"

/db_xref="GI:187152"

/db_xref="GDB:G00-120-153"

/translation="MKMPLGVCLVLPVHSEGGGKLTAVDPETNNVSEIISY
GPPSEYLVETEDGYILCLNRIPIHGRKHSDKGPVDFLOHGLLADSSNWTNLIANS
SIGFILLADGFDVWNGNRSNTWRKHLSVSQDEFWAFSDIDEAKYDLPASIMFIL
NKTGQEVYVYVHSGQTIGFTAFSQIPELAKRIKMFALGPVAVSFAVCTSPAKLGR
LDPLHILFDGKEFLPQSAFLKVLGTHTVILKELCGNLCTLCCGFNERNLNMS
VDVYITDTPAGTSVQNMHWSQAVKQKFAQFDGSSAKNYFHYNQSPPTVYNNKDM
VPTAVWSGGHDLADVVDVNIILQTITNLVPHTPEWEHLDFINGLDPAPWRLYNNKI
NLMRXQO"

41..103

/sig_peptide

/gene="LIPA"

/note="alternate signal peptide may begin at position 47;
as predicted from von Heijne, G. (1986) Nucl. Acids Res.
14,4683-4690; G00-120-153"

104..1237

/gene="LIPA"

/product="lysosomal acid lipase/cholesterol esterase"

/EC_number="3.1.1.13"

/note="G00-120-153"

2466..2471

/polyA_signal

```

/genes="LIPA"
/notes="G00-120-153"
2493
polya_site
/genes="LIPA"
/notes="G00-120-153"
692 a 484 c 516 g 801 t
BASE COUNT
ORIGIN
Query Match 25.7%; Score 526.2; DB 9; Length 2493;
Best Local Similarity 56.6%; Pred. No. 7.1e-124;
Matches 751; Conservative 0; Mismatches 376; Indels 0; Gaps 0;
QY 240 AAGCTGTGTGACCCAGAGCAATTCATGAATATTAGTGAATATCATCAACATCAAGCTAT 299
Db 119 ACAGCTGTGATCCTGAAACAAACATCAATGTGAGTGAATATCTCTTACTGGGATTC 178
QY 300 CCCTGTGAGGAATATGAGTCGCACTGAGTGGGTATATCTCTTCTTAACAGGATT 359
Db 179 CTAAGTGAGGAATACCTAGTTGAGACAGAGATGGATATATCTGTGCTTAAACGAATT 238
QY 360 CCTCGAGGCTAGTCAACCTAAGAACAGACAGGTTCCAGGCTGTGTGTGTACTGACGAT 419
Db 239 CCTCATGGGAGGAACCACTCTGACAAAGGTCCCAACCAAGTTGTCTTCTGCAACAT 298
QY 420 GGCTAGTTGGAGTGTCTAGCACTGGATTCCAACTGCTGCCCAACATAGCTGGCTTC 479
Db 299 GGCTTGTGCGAATCTTAGTAACCTGGTCCAAACCTTGCACACAGCAGCTGGCTTC 358
QY 480 ATTCTGCGCATGTGCTGTTTGTACGTGTGGATGGGAAACAGACAGGGGAAACGCTGGTCT 539
Db 359 ATCTTGTCTGCTGTTTGTACGTGTGGATGGGCAACAGCAGAGGAAATACCTGGTCT 418
QY 540 CGAAACACAGACACTCTCCATAGACCAAGATGAGTCTGGGCTTTTCAGTTATGATGAG 599
Db 419 CGAAACACATAGACACTCTCAGTTCTCAGATGAATTTCTGGGCTTTTCAGTTATGATGAG 478
QY 600 ATGCTAGTTTGTACCTCTCTGAGTGATATAAATTATTTTGCAGAAACGGGCCAGGAA 659
Db 479 ATGCAAAATATGACCTACCACTGCTTCAATTTCACTTGAATAAATCTGCCAAGAA 538
QY 660 AAGATCTATTATGTCGGCTATTCAAGGCACCAACAGGCTTTATTCAGTTTCCACC 719
Db 539 CAAGTGATTATGTGGTCTATCTCAAGGCACCACTATAGTTTATAGCAATTTTCCAG 598
QY 720 ATGCCAGAGTGGTCAAGAAATCAAAATGATTTTGTCTTACACCCATAGCACTGTT 779
Db 599 ATCCCTGAGCTGGCTAAAGGATTAATAATGTTTTTGCCTGGGCTCTGGCTTCCGTC 658
QY 780 AAGCATGCAAAAGCCCGGACCAAAATTTTGTGTGCTGCAGATATGATGATCAAGGGA 839
Db 659 GCCTTCTGTACTAGCCCTATGGCCAAATTAGGACGATTACCAATCATCTCAATTAAGGAC 718
QY 840 TTGTTGGCAAAAGAAATTTCTGTATCAGACAGATTTCTCAGACAACTGTGTTATTAC 899
Db 719 TTATTTGGAGCAAAAGAAATTTCTTCCACAGATGCGTTTTTGAAGTGGCTGGGTACCCAC 778
QY 900 CTTTGTGCGCAGGATTTCTTGATCAGATTTGTAGTAATATCATGTTACTTCTGGGTGGA 959
Db 779 GTTTGCACTATGTCATCTAGGAGGACTCTGTGGAAATCTCTGTTTCTTCTGTGGA 838
QY 960 TTCACACCAACAAATATGAACATAGCGCGAGCAAGTGTATATGCTGCCACACTCTTGCT 1019
Db 839 TTTAATGAGAGAAATTTAAATATGCTAGAGTGGATGTATATACACACATTTCTCTGCT 898
QY 1020 GGCACATCTGCAAAATTTCTACATGCGGACCGGACGAGTGTATCTGGTGAATCTCCGG 1079
Db 899 GGAATCTCTGTGCAAAACATGTTACATGGAGCCGCTGTAAATTTCCAAAAGTTTCAA 958
QY 1080 GCATTTGACTGGGGAGTGAGACCAAAATCTGAAAATTCGAATCAATCAGCAACTCTGTGA 1139
Db 959 GCCTTTGACTGGGAGGACGAGTCCCAAGATTTATTTTCAATACACAGAGTATCTCTCC 1018
QY 1140 AGGTACAGAGTCAGAGATATGACGCTCCCTACAGCAATGTGGACAGAGGTCAAGACTGG 1199

```

```

Db 1019 ACATCAATGTGAAGGACATGCTTGTGCCACTGCACTGTGGAGCGGGGTCAAGACTGG 1078
QY 1200 CTTTCAATCCAGAGAGCGTGAATGCTCTCTCTGAGTGACCAACCTCATCTACCAT 1259
Db 1079 CTTGAGATGCTACGACGTCATATCTTACTGACTCAGATCACCACCTGGTGTCCAT 1138
QY 1260 AAGAATATCTCTGAATGGGCTCAGTGGATTTTCATCTGGGGTTTGATGCTCTCTACCT 1319
Db 1139 GAGAGCATTCGGAATGGGAGCATCTTGACTTCATTTGGGGCTGGATGCCCTTGGAGG 1198
QY 1320 ATGTCAATCAATCATCTGATGTCAGCAGGAGAGACCACT 1366
Db 1199 CTTTATAATAAATTATTAATCTAATGAGGAATATCAGTGAAGCT 1245

RESULT 15
BC012287 2586 bp mRNA linear PRI 08-AUG-2001
LOCUS Homo sapiens, similar to lipase A, lysosomal acid, cholesterol
DEFINITION esterase (Wolman disease), clone MGC:5229 IMAGE:2900168, mRNA,
complete cds.
ACCESSION BC012287 GI:15126726
VERSION BC012287.1
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2586)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (08-AUG-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: gcapsb-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: villalob@bcm.tmc.edu.
Villalob, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAP Plate: 3 Row: c Column: 3
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 434305.
Location/Qualifiers
1. 2586
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:5229 IMAGE:2900168"
/tissue_type="Placenta, choriocarcinoma"
/clone_lib="NIH_MGC_10"
/lab_host="DH10B"
/note="vector: pCMV-SPORT6"
101..1300
/codon_start=1
/product="Similar to lipase A, lysosomal acid, cholesterol
esterase (Wolman disease)"
/protein_id="AAH12287.1"
/db_xref="GI:15126726"
/translation="MKMRFLGLVCLVPLHSEGGKLTALDPETNNVSEIISYM"

FEATURES
source
CDS

```

GPPSEYLVETEDGVILCLNRI PHGRKNHSDKPKVPVFLQHLGLADSNVNTLANS
SLGFLADAGDVMSRSGNTWRSKTKLTVSODEFWAFSPYDEMAKYDLPASINPL
NKTOBOVYVYHSGQTTIGFIAEQIPELAKRIKMFALGPVAFCTSPMAKLR
LPDLHLDKDFELFQSAFLKWLGHVCHVILKELCNELFCFLCGFNERLNMSR
VDTITSPAGTSVONMLHWSQAVKFORQAFDWGSSAKNYFHYNQSYPTYNVKOML
VPTAVMSGGHMDLADVVDVNVILLQITNLVPHESIPEWEHLDFIWLGLDAPWRLYNKI
NLMRYQ

BASE COUNT 736 a 511 c 531 g 808 t

Query Match 25.7%; Score 524.6; DB 9; Length 2586;
Best Local Similarity 66.5%; Pred. No. 1.8e-137;
Matches 750; Conservative 0; Mismatches 327; Indels 0; Gaps 0;

QY	240	AAAGCTGTGGACCCAGAGCAATTCATGAATATTAGTGAATATCATCAACATCAAGGCTAT	299
DB	179	ACAGCTTTGGATCCTGAACAAACATGAATGTGAGTGAATATCTCTTACTGGGGATTC	238
QY	300	CCCTGTGAGGAATATGAAGTGCBACTGAAGATGGGTATATCCCTTCTGTGTAAAGGATT	359
DB	239	CCTAGTGAGGAATACCTAGTTGAGACAGATGGATATATCTGTGCTTAAACCGAATT	298
QY	360	CCTGAGGCTTAGTGAACCTTAAGAGACAGTTCAGGCTGCTGCTGCTGCTGCTGCTGCT	419
DB	299	CCTCATGGAGGAGAACCATCTGCAAAAGTCCCAACAGTGTCTTCTCTGCAACAT	358
QY	420	GGCCTAGTTGGAGTGTAGCAACTGGATTTCCAACTGCCCAACAATAGCCTGGGCTTC	479
DB	359	GGCTTGCTGGCAGATCTTAGTAACTGGGTACAAACCTTGCCACAGCAGCCTGGGCTTC	418
QY	480	ATCTGGCAGATGCTGTTTGAAGTGTGATGGGACAGCAGGGAACGCTGCTCT	539
DB	419	ATCTTGCTGATGCTGTTTGAAGTGTGATGGGACAGCAGGGAATACCTGCTCT	478
QY	540	CGAAACCAAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTTCAGTTATGATGAG	599
DB	479	CGGAACATAAGACACTCTCAGTTTCTCAGGATGAATCTGGGCTTTTCAGTTATGATGAG	538
QY	600	ATGGCTAGGTTGACCTTCCTGAGTGAATAACTTTATTTGACAGAAACCGGCCAGAA	659
DB	539	ATGGCAAAATATGACCTTACCAGCTTCCATTAACCTTCTGTAATAAACTGGCCAAAG	598
QY	660	AAGATCTATTATGCTGGCTATTACAGGGCACCACCATGGCTTTTATGCAATTTCCACC	719
DB	599	CAAGTGTATTATGCTGGCTATTCTCAGGCACCACTATAGTTTATAGCAATTTTCAG	658
QY	720	ATGCCAGAGCTGGCTCAGAAATCAAAATGATTTTGTCTTTAGCACCCTATAGCCACTGTT	779
DB	659	ATCCCTGAGCTGGCTTAAAGGATTAATGTTTGTGCTGGCTGCTGCTGCTGCTGCTGCT	718
QY	780	AAGCATGCAAAAGCCCGGACCAAAATTTTGTGCTGCCAGATATGATGATCAAGGA	839
DB	719	GCCTTCTGTACTAGCCCTATGGCCAAATTAGGACGATTACCAAGATCATCTCAATTAGGAC	778
QY	840	TTGTTTGGCAAAAGAAATTTCTGATCAGACCAAGATTTCTCAGACAACTGTTTATTATAC	899
DB	779	TTATTTGGAGACAAGAAATTTCTCCAGAGTGGCTTTTGAAGTGGCTGGTACCCAC	838
QY	900	CTTTTGGCCAGGTAATCTTGATCAGATTTGTAGTAATATCATGTTACTTCTGGGTGA	959
DB	839	GTTTGCACTCATGTACTAGGAGGCTCTGTGAAATCTCTGTTTCTTCTCTGTGTGA	898
QY	960	TTCAACACCAAAATATGAATGAGCGGAGTGTATATGCTGCCACACTCTTGTCT	1019
DB	899	TTTATGAGAGAAATTTAAATATGTCTAGAGTGGATGATATACAAACATCTCTCTGCT	958
QY	1020	GGAACTCTGTGCAAAATATTCTACCTGGAGCCAGGAGTGAATTTCTGTGAACCTCGG	1079
DB	959	GGAACTTCTGTGCAAAACATGTTACCTGGAGCAGGCTGTTAAATCCAAAAGTTTCAA	1018
QY	1080	GCATTTGCTGGGAGTGAACCAAAATCTGGAAAATGCAATCAGCCAACTCTCTGTA	1139
DB	1019	GCCTTTGCTGGGGAAGCAGTGCCAAAGAAATTTTTCATTACACCAAGATTATCTCTCC	1078

Search completed: February 19, 2004, 08:29:14
Job time : 7525.62 secs

QY	1140	AGGTACAGATCAGAGATATGACGCTCCCTACAGCAATGTGACAGGAGGTACAGACTGG	1199
DB	1079	ACATCAATGTGAAGACATGCTTGTGCCGACTGCAGTCTTGAGCGGGGTACAGACTGG	1138
QY	1200	CTTTCAAATCCAGAGACGTGAAAATGCTGCTCTCTGAGGTGACCAACCTCATCTTACCAT	1259
DB	1139	CTTGCAATGTCTACGACGTCAATATCTTACTGACTCAGATCACCACCTTGGTGTTCAT	1198
QY	1260	AAGATATTCTGATGGCTCACTGATGATTTTCACTGGGTTTGGATGCTCTCACCGT	1319
DB	1199	GAGAGCATTTCCGAAATGGAGCATCTTGACTTCAATTGGGGCTGGATGCCCTTGGAGG	1258
QY	1320	ATGTACAATGAATCATCCATCTGATGACAGGAGGAGACCAACCT	1366
DB	1259	CITTTAATAAATTTAATTAATCTAATGAGAAATATCAGTGAAGCT	1305

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 04:41:17 ; Search time 563.905 Seconds
(without alignments)
9784.725 Million cell updates/sec

Title: US-10-042-431-45
Perfect score: 2044
Sequence: 1 gtcaccacacgctcgggg.....aaaaaaagggcgccgc 2044

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03;+
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT: *
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT: *
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT: *
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT: *
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT: *
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT: *
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT: *
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT: *
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT: *
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT: *
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT: *
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT: *
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT: *
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT: *
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT: *
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT: *
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT: *
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT: *
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT: *
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT: *
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT: *
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT: *
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT: *
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT: *
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2044	100.0	2044	22 AAF45131	Human TANGO 294 cD
2	2044	100.0	2044	25 ABK94106	cDNA encoding huma
3	1993.8	97.5	2035	22 AAD19226	Human CG162 (or C5
4	1343.4	65.7	1384	22 AAD19220	Human CG162 (or C5
5	1269	62.1	1269	22 AAF45132	Human TANGO 294 OR
6	1138	55.7	1210	24 ABK96569	Human cDNA encodin
7	1126	55.1	1206	24 ABS55356	cDNA encoding huma
8	1107.8	54.2	1272	24 ABK90871	cDNA encoding huma

9	890.4	43.6	18554	24 ABS55357	Genomic DNA encodi
10	752	36.8	1116	24 ABK96570	Human cDNA encodin
11	527.8	25.8	2481	24 ABK97363	Gene #3881 used to
12	526.2	25.7	2493	22 AAH02906	Human shear stress
13	489.2	23.9	1378	14 AAQ42310	RGL precursor. Or
14	470	23.0	1367	7 AAN60685	Sequence encoding
15	470	23.0	1367	7 AAN60566	Sequence encoding
16	470	23.0	1367	17 AAT58916	Human gastric lipa
17	468.4	22.9	1695	22 AAH57420	Human stomach cell
18	464.4	22.7	1336	6 AAN50385	Rat prelingual lip
19	456.8	22.3	1411	24 ABL57144	Kid goat pregastr
20	454.2	22.2	1194	24 AAD27803	Human lysosomal ac
21	454	22.2	1228	24 ABK12385	cDNA encoding huma
22	450.8	22.1	1284	24 ABK85978	DNA encoding human
23	449.8	22.0	8324	24 ABL57145	Yeast YB-1 express
24	448.4	21.9	1473	24 ABK12388	cDNA encoding huma
25	446.8	21.9	1185	24 AAD31193	Human triacylglyce
26	444.2	21.7	1528	15 AAQ68388	Canine gastric lip
27	444.2	21.7	1528	17 AAT58915	Dog gastric lipase
28	444.2	21.7	1531	15 AAQ68389	Canine gastric lip
29	443.6	21.7	1200	22 AAH78205	Nucleotide sequenc
30	424.8	20.8	1260	24 ABS71703	DNA encoding human
31	419.8	20.5	1086	25 ACC44783	Human triacylglyce
32	414.2	20.3	1360	24 AAL47480	Human lysosomal ac
33	414.2	20.3	1360	25 ABK14880	Human lipase cDNA.
34	374.4	18.3	996	25 ACC46563	Human dithp secret
35	346.4	16.9	1138	24 ABQ82326	Human NOV1 encodin
36	332.8	16.3	1098	24 AAD27802	Human lysosomal ac
37	319.4	15.6	1143	24 ABK12386	cDNA encoding huma
38	315.4	15.4	828	24 AAD31194	Human triacylglyce
39	300.4	14.7	1377	24 ABS71702	DNA encoding human
40	290	14.2	842	24 AAD31192	Human triacylglyce
41	286	14.0	699	22 AAF28679	Human protein HP03
42	286	14.0	1308	22 AAF28689	Human protein HP03
43	282.8	13.8	802	22 AAI97406	Human neuroblastom
44	257.8	12.6	617	21 AAA44349	Human secreted exp
45	244.6	12.0	1971	23 AAS91026	DNA encoding novel

ALIGNMENTS

RESULT 1
AAF45131
ID AAF45131 standard; cDNA; 2044 BP.
XX
XX AAF45131;
XX
XX
DT 30-MAR-2001 (first entry)
XX
DE Human TANGO 294 cDNA.

XX Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
XX central nervous system; focal brain disorder; bipolar affective disorder;
KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
KW neuropsychiatric; psychoactive substance use; anxiety; ss.
XX
OS Homo sapiens.

XX WO2000077239-A2.

XX 21-DEC-2000.

XX 24-MAY-2000; 2000WO-US14858.

XX 14-JUN-1999, 99US-0333159.

XX (MILL-) MILLENNIUM PHARM INC.

XX McCarthy SA, Fraser CC, Sharp JD, Barnes TM;

DR	WPI; 2001-032313/04.	QY	661	AGATCTATTGTCGGCTATTTCACAGGCGACCACTGGGCTTTATTGCACTTTTCCACCA	720
DR	P-PSDB; AAB66045.	DB	661	AGATCTATTGTCGGCTATTTCACAGGCGACCACTGGGCTTTATTGCACTTTTCCACCA	720
XX		QY	721	TGCCAGAGCTGGCTCAGAAAAAATAAATGATTTGCTTTAGCACCACCATGCACTGTTA	780
PT	TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for	DB	721	TGCCAGAGCTGGCTCAGAAAAAATAAATGATTTGCTTTAGCACCACCATGCACTGTTA	780
PT	screening assays and diagnostic assays and for the treatment of	QY	781	AGCATGCAAAAAGCCCGGACCAAAATTTTGTGCTGCCAGATATGATCATCAAGGAT	840
PT	neurological diseases such as Alzheimer's, Parkinson's and Huntington's	DB	781	AGCATGCAAAAAGCCCGGACCAAAATTTTGTGCTGCCAGATATGATCATCAAGGAT	840
PT	disease -	QY	841	TGTTGGCAAAAAGAAATTTCTGTATCAGACAGCAATTTCTCAGACAACTTTGTTATTACC	900
XX	Claim 1; Fig 6; 359pp; English.	DB	841	TGTTGGCAAAAAGAAATTTCTGTATCAGACAGCAATTTCTCAGACAACTTTGTTATTACC	900
CC	The present invention relates to TANGO or INTERCEPT proteins and coding	QY	901	TTTGTGGCCAGGTGATTTCTGATCAGATTTGTAGTATATATCATGTTCTCTGGGTGAT	960
CC	sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,	DB	901	TTTGTGGCCAGGTGATTTCTGATCAGATTTGTAGTATATATCATGTTCTCTGGGTGAT	960
CC	sequences are useful for the treatment of neurological disorders and coding	QY	961	TCAACACCAACAATATGAAATGAGCCGAGCAAGTGTATATGCTGCCACCACTTTCTGCTG	1020
CC	central nervous system (CNS) disorders, CNS-related disorders, focal	DB	961	TCAACACCAACAATATGAAATGAGCCGAGCAAGTGTATATGCTGCCACCACTTTCTGCTG	1020
CC	brain disorders, global-diffuse cerebral disorders and other	QY	1021	GAAATCTGTGCAAAAATATTTCTACACTGGAGCCAGGAGTGAATTTCTGGTGAATCTCCGG	1080
CC	neurological and cerebrovascular disorders. The CNS disorders include	DB	1021	GAAATCTGTGCAAAAATATTTCTACACTGGAGCCAGGAGTGAATTTCTGGTGAATCTCCGG	1080
CC	Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic	QY	1081	CATTGTGCTGGGAGTGAGACCAAAAATCTGGAATAATGCAATCAGCAACTCTCTAA	1140
CC	lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,	DB	1081	CATTGTGCTGGGAGTGAGACCAAAAATCTGGAATAATGCAATCAGCAACTCTCTAA	1140
CC	autonomic function disorders such as hypertension and sleep disorders,	QY	1141	GCTCAGAGTCAGAGATATGACGGTCCCTACAGCAATGTGGACAGGAGTCAAGACTGGC	1200
CC	neuropsychiatric disorders, psychoactive substance use disorders,	DB	1141	GCTCAGAGTCAGAGATATGACGGTCCCTACAGCAATGTGGACAGGAGTCAAGACTGGC	1200
CC	anxiety, and bipolar affective disorder.	QY	1201	TTTCAATCCAGAGACGTGAAATGCTCTCTGAGGTGACCAACTCATCTACCATA	1260
XX	Sequence 2044 BP; 618 A; 401 C; 460 G; 565 T; 0 other;	DB	1201	TTTCAATCCAGAGACGTGAAATGCTCTCTGAGGTGACCAACTCATCTACCATA	1260
QY	Query Match 100.0%; Score 2044; DB 22; Length 2044;	QY	1261	AGAAATATCTCGAATGGGCTCACTGGGTTTTCATCTGGGTTTGGATGCTCTCACCGTA	1320
DB	Best Local Similarity 100.0%; Pred. No. 0;	DB	1261	AGAAATATCTCGAATGGGCTCACTGGGTTTTCATCTGGGTTTGGATGCTCTCACCGTA	1320
DB	Matches 2044; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	QY	1321	TGTAATATGAAATCATCTGATGACAGGAGGAGACCAACTTTTCCAGGAGCGGT	1380
QY	1 GTCACCCAGCGTCCGGGAATGTCAGAGGAAATATGTAAGAGTTTAAACCCAC	DB	1321	TGTAATATGAAATCATCTGATGACAGGAGGAGACCAACTTTTCCAGGAGCGGT	1380
DB	1 GTCACCCAGCGTCCGGGAATGTCAGAGGAAATATGTAAGAGTTTAAACCCAC	QY	1381	GTGAGGCGGTATTGTGAAGCATCTGACACTGACATCTTAGGACAACTCTCTGAGGATG	1440
QY	61 AAATCTCTTACTTTAGAAATGATGTTTCCAGAGAAATGTAATTCAGTACATATGCCACTA	DB	1381	GTGAGGCGGTATTGTGAAGCATCTGACACTGACATCTTAGGACAACTCTCTGAGGATG	1440
DB	61 AAATCTCTTACTTTAGAAATGATGTTTCCAGAGAAATGTAATTCAGTACATATGCCACTA	QY	1441	GGGCTAGAACCCATGAAGGCAAAATTAAGGAGAGAGAGAGCTAGTATACATTTTTCAGA	1500
QY	121 GGACCATGTTGGAACCTTGTCAAGACAGTGGATTTCTCAGACAGATGGAATGTGGC	DB	1441	GGGCTAGAACCCATGAAGGCAAAATTAAGGAGAGAGAGAGCTAGTATACATTTTTCAGA	1500
DB	121 GGACCATGTTGGAACCTTGTCAAGACAGTGGATTTCTCAGACAGATGGAATGTGGC	QY	1501	TTCCCTGACACTTGGCACTAAATCCGACACTTACATTTTCTGTAATAATAA	1560
QY	181 TTCTGATTCCTGCTGGCTATATGTTCCAGAGAAATGTAATTCAGTACATATGCCACTA	DB	1501	TTCCCTGACACTTGGCACTAAATCCGACACTTACATTTTCTGTAATAATAA	1560
DB	181 TTCTGATTCCTGCTGGCTATATGTTCCAGAGAAATGTAATTCAGTACATATGCCACTA	QY	1561	GTACTTATTTAGTAAATAGAGGTTTGTATGCTTATATATTTTACCATCTTGAAGGAT	1620
QY	241 AAGCTGTGGACCCAGAGACATTCATGAATATTAGTGAATATCCAAATCAAGGCTATC	DB	1561	GTACTTATTTAGTAAATAGAGGTTTGTATGCTTATATATTTTACCATCTTGAAGGAT	1620
DB	241 AAGCTGTGGACCCAGAGACATTCATGAATATTAGTGAATATCCAAATCAAGGCTATC	QY	1621	AGGTTTTTACCTGATAGCAGAAATATCTAGACATCTCTATATCATTCAGTAAATCTC	1680
QY	301 CCTGTGAGGATATGAGTCCCACTGAGATGGGTATATCCTTCTGTTAAACAGGATTC	DB	1621	AGGTTTTTACCTGATAGCAGAAATATCTAGACATCTCTATATCATTCAGTAAATCTC	1680
DB	301 CCTGTGAGGATATGAGTCCCACTGAGATGGGTATATCCTTCTGTTAAACAGGATTC	QY	1681	TTTAAAAACACCTTATTTGTTTCTTAAAGCCTATTTTGGAGCACTAAAGTAAATGGC	1740
QY	361 CTCAGAGGCTAGTCAACCTTAAGAGACAGGTTCCAGGCTGTGGTGTACTGCAGCATG	DB	1681	TTTAAAAACACCTTATTTGTTTCTTAAAGCCTATTTTGGAGCACTAAAGTAAATGGC	1740
DB	361 CTCAGAGGCTAGTCAACCTTAAGAGACAGGTTCCAGGCTGTGGTGTACTGCAGCATG	QY	1741	AAATGGGACAGATATGAGGCTCTGAGTCTGTGGATTTATTTGTTGACATTTTGAACAAATAA	1800
QY	421 GCCTAGTTGGAGGTGCTAGAACTGGATTTCCAACTGCCCCCAACATAGCCTGGCTTCA	DB			
DB	421 GCCTAGTTGGAGGTGCTAGAACTGGATTTCCAACTGCCCCCAACATAGCCTGGCTTCA	QY			
QY	481 TTCTGGCAGATGCTGTTTTCAGCTGTGGATGGGGAACAGCAGGGGAAACGCTGTGCTC	DB			
DB	481 TTCTGGCAGATGCTGTTTTCAGCTGTGGATGGGGAACAGCAGGGGAAACGCTGTGCTC	QY			
QY	541 GAAACACACAGACATCTCCATAGACCAAGATGATTTCTGGGCTTTTCACTGATGAGA	DB			
DB	541 GAAACACACAGACATCTCCATAGACCAAGATGATTTCTGGGCTTTTCACTGATGAGA	QY			
QY	601 TGGCTAGGTTTGAACCTTCTCCAGTGTAACTTTATTTTGCAGAAACCGGCGCAGGAA	DB			
DB	601 TGGCTAGGTTTGAACCTTCTCCAGTGTAACTTTATTTTGCAGAAACCGGCGCAGGAA				

Db 481 TTTCTGGCAGATGCTGTTTGAAGTGGATGGGAAACAGCAGCGGAAACGCGTGTCTC 540
 QY 541 GAAACACAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTTCAGTTATGATGAGA 600
 Db 541 GAAACACAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTTCAGTTATGATGAGA 600
 QY 601 TGGCTAGGTTTGAACCTTCTGCAAGTATGATAAATTTTGTGAGAGAAACGGGCGCAGGAAA 660
 Db 601 TGGCTAGGTTTGAACCTTCTGCAAGTATGATAAATTTTGTGAGAGAAACGGGCGCAGGAAA 660
 QY 661 AGATCTATATATGCGGCTATTCACAGGGCACCACCATGCGCTTTATGCAATTTCCACCA 720
 Db 661 AGATCTATATATGCGGCTATTCACAGGGCACCACCATGCGCTTTATGCAATTTCCACCA 720
 QY 721 TGGCAGAGCTGGCTCAGAAATCAAAATGTAATTTTGTGAGCACCACCATGAGCCTGTTA 780
 Db 721 TGGCAGAGCTGGCTCAGAAATCAAAATGTAATTTTGTGAGCACCACCATGAGCCTGTTA 780
 QY 781 AGCATGCAAAAGCCCGGACCAAAATTTTGTGAGCACCACCATGAGCCTGTTA 840
 Db 781 AGCATGCAAAAGCCCGGACCAAAATTTTGTGAGCACCACCATGAGCCTGTTA 840
 QY 841 TGGTTGCAAAAGAAATTTCTGATCAGACCAAGATTTCTCAGACAACTTTGTTATTTACC 900
 Db 841 TGGTTGCAAAAGAAATTTCTGATCAGACCAAGATTTCTCAGACAACTTTGTTATTTACC 900
 QY 901 TTTGTGCGCAGGTGATTTCTGATCAGATTTGTAATATATCATTTCTGGGTGGAT 960
 Db 901 TTTGTGCGCAGGTGATTTCTGATCAGATTTGTAATATATCATTTCTGGGTGGAT 960
 QY 961 TCAACACCAATATGAAATGAGCGGACGAGTGTATATGTCGCCACACTTTGCTG 1020
 Db 961 TCAACACCAATATGAAATGAGCGGACGAGTGTATATGTCGCCACACTTTGCTG 1020
 QY 1021 GAACATCTGTGCAAAATATTTCTACCTGGAGCCAGGAGTGAATTTCTGTGAACCTCCGG 1080
 Db 1021 GAACATCTGTGCAAAATATTTCTACCTGGAGCCAGGAGTGAATTTCTGTGAACCTCCGG 1080
 QY 1081 CATTTGATGGGGAGTGAGACCAAAATCTGGAATAATGCAATCAAGCCTCTGTAA 1140
 Db 1081 CATTTGATGGGGAGTGAGACCAAAATCTGGAATAATGCAATCAAGCCTCTGTAA 1140
 QY 1141 GGTACAGAGTCAGAGATATGAGGTCCTACAGCAATGTGACAGAGGTCAGGCTGC 1200
 Db 1141 GGTACAGAGTCAGAGATATGAGGTCCTACAGCAATGTGACAGAGGTCAGGCTGC 1200
 QY 1201 TTTCAAAATCCAGAGAGCTGAAATGCTGCTCTGAGGTGACCAACCTCATCTACATA 1260
 Db 1201 TTTCAAAATCCAGAGAGCTGAAATGCTGCTCTGAGGTGACCAACCTCATCTACATA 1260
 QY 1261 AGAATATTTCTGAATGGGCTCAGTGGATTTCTGCGGTTTGGATGCTCTCCCTCAGCTA 1320
 Db 1261 AGAATATTTCTGAATGGGCTCAGTGGATTTCTGCGGTTTGGATGCTCTCCCTCAGCTA 1320
 QY 1321 TGTACAAATGAATCATCCATCTGATGAGCAGGAGGAGACCAACCTTTCCAGGGACGCT 1380
 Db 1321 TGTACAAATGAATCATCCATCTGATGAGCAGGAGGAGACCAACCTTTCCAGGGACGCT 1380
 QY 1381 GTGAGGCGGTATTTGAGCATCTGACATGACATCTTAGGACCAACCTCTGAGGGATG 1440
 Db 1381 GTGAGGCGGTATTTGAGCATCTGACATGACATCTTAGGACCAACCTCTGAGGGATG 1440
 QY 1441 GGGCTAGGACCCATGAAGGAGCAATACGAGAGCAGAGACCTTAGTATACATTTTTCAGA 1500
 Db 1441 GGGCTAGGACCCATGAAGGAGCAATACGAGAGCAGAGACCTTAGTATACATTTTTCAGA 1500
 QY 1501 TTCCCTGCACTGGCACTAAATCCGACATTAATTACATTTTCTGTAAATTTAA 1560
 Db 1501 TTCCCTGCACTGGCACTAAATCCGACATTAATTACATTTTCTGTAAATTTAA 1560
 QY 1561 GTACTTATAGGTAATAGAGGTTTGTGATGCTTATATATTTCTACCATCTTGAAGGT 1620

RESULT 3

AAD19226
 ID AAD19226 standard; DNA; 2035 BP.

XX AAD19226;

XX 18-DEC-2001 (first entry)

XX Human CG162 (or C59) lipase DNA #2.

Human; apolipoprotein; lipase; lipoprotein receptor; ALLr; angina; cardiovascular disease; lipid metabolism; myocardial infarction; cerebral ischaemia; arterial thrombosis; thrombolytic; antilipemic; coronary artery thrombosis; cerebral artery thrombosis; stroke; intracardiac thrombosis; gene therapy; cardiovascular; vasodilator; neuroprotectant; cerebroprotective; ds.

XX Homo sapiens.

Key Location/Qualifiers

FT CDS 189..1418
 /*tag= a
 /product= "Human CG162 (or C59) lipase protein"

FT sig_peptide 189..245

FT mat_peptide 246..1415

FT /*tag= c
 /product= "Human mature CG162 (or C59) lipase protein"

XX WO200179446-A2.

XX 25-OCT-2001.

XX 16-APR-2001; 2001WO-US12529.

XX 14-APR-2000; 2000US-197137P.

XX 20-JUN-2000; 2000US-0598042.

PR 03-AUG-2000; 2000US-0631451.
 PR 22-SEP-2000; 2000US-0667298.
 PR 17-NOV-2000; 2000US-0714936.
 PA (HYSE-) HYSEQ INC.
 XX Ballinger DG, Loeb D, Montgomery JR, Tang TY, Zhou P, Goodrich R;
 PI Liu C, Asundi V, Zhao QA, Wehrman JT, Drmanac RT, Ren F, Qian XB;
 PI Wang D;
 XX WPI: 2001-611724/70.
 DR P-PSDB; AAE11931.
 XX Nucleic acids encoding human apolipoproteins, lipases, and lipoprotein
 PT receptor polypeptides, useful for preventing diagnosing and treating
 PT lipid metabolism disorders, thrombosis and cardiovascular diseases -
 XX Claim 1; Page 183-185; 266pp; English.
 XX The invention relates to polynucleotides encoding proteins CG122, CG179,
 CC CG95, CG121, CG162, CG27, CG153 and CG168 which are related to proteins
 CC involved in lipid metabolism and cardiovascular disease such as human
 CC apolipoproteins, lipases and lipoprotein receptor proteins. These DNA
 CC and protein sequences are useful for treating or preventing disorders
 CC associated with apolipoproteins, lipases and lipoprotein receptor (ALLR)
 CC expression and for treating lipid metabolism, cardiovascular diseases
 CC and thrombosis. Antibodies against these proteins are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of these sequences. ALLR polypeptides are also
 CC useful for identifying agents (agonists and antagonists) that bind to
 CC them and cells expressing ALLR proteins are useful for identifying a
 CC therapeutic agent for use in treatment of a pathology related to
 CC aberrant expression or physiological interactions of this polypeptide.
 CC Vectors comprising these DNA and protein sequences are also useful for
 CC producing ALLR proteins. The nucleic acids and polypeptides of the
 CC invention are also useful for the treatment of occlusive cardiovascular
 CC diseases, myocardial infarction, cerebral ischaemia, angina, arterial
 CC thrombosis, coronary artery thrombosis and cerebral artery thrombosis
 CC or intracardiac thrombosis and stroke. The nucleotides of the invention
 CC are used in gene therapy. The present sequence is human CG162 (or C59)
 CC lipase DNA.
 XX
 SQ Sequence 2035 BP; 608 A; 395 C; 456 G; 576 T; 0 other;
 Query Match 97.5%; Score 1993.8; DB 22; Length 2035;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1995; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 18 GGGAAATTCAGCAGGAAATATGTGAAGATTTTAAACCCCAAAATTCCTTACTTTA 77
 DB 39 GGGAAATTCAGCAGGAAATATGTGAAGATTTTAAACCCCAAAATTCCTTACTTTA 98
 QY 78 GAATTAAGTTTACATTCGAGGAAATTAATGATGAGATTTGGAACATTTGGAAC 137
 DB 99 GAATTAAGTTTACATTCGAGGAAATTAATGATGAGATTTGGAACATTTGGAAC 158
 QY 138 TTGTCAAGACAGTGGATGTCACACAGAAATGGAATGAGCTTCTGATTTCTGGTGG 197
 DB 159 TTGTCAAGACAGTGGATGTCACACAGAAATGGAATGAGCTTCTGATTTCTGGTGG 218
 QY 198 TATATGTTCCAGAGAAATGTGAATTCAGTACATATGCCAACTAAAGCTGTGGACCCAGAA 257
 DB 219 TATATGTTCCAGAGAAATGTGAATTCAGTACATATGCCAACTAAAGCTGTGGACCCAGAA 278
 QY 258 GCAATTCATGATATTAAGTGAATATCAATCAATCAAGGCTATCCCTGTGAGGATATGAA 317
 DB 279 GCAATTCATGATATTAAGTGAATATCAATCAATCAAGGCTATCCCTGTGAGGATATGAA 338
 QY 318 GTGCAACTGAAGATGGGTATATCTTTCTGTGAACAGATTCCTCGAGGCTTAGTGCAA 377
 DB 339 GTGCAACTGAAGATGGGTATATCTTTCTGTGAACAGATTCCTCGAGGCTTAGTGCAA 398
 QY 378 CCTAAGAAGACAGGTTCCAGGCTGTGGTGTACTGACAGATGGCCCTAGTTGGAGGTGCT 437

DB 399 CCTAAGAAGACAGGTTCCAGGCTGTGGTGTACTGACAGATGGCTAGTTGGAGGTGCT 458
 QY 438 AGCAACTGGATTTCCAACTGCGCCCAATAGCTGGGCTTCATTCGCGCATGCTGTGT 497
 DB 459 AGCAACTGGATTTCCAACTGCGCCCAATAGCTGGGCTTCATTCGCGCATGCTGTGT 518
 QY 498 TTTGACGTTGGATGGGGAACAGCAGGGAACAGCTGCTCGAATAACACAGACATCTC 557
 DB 519 TTTGACGTTGGATGGGGAACAGCAGGGAACAGCTGCTCGAATAACACAGACATCTC 578
 QY 558 TCCATAGACCAAGATGATGTTCTGGGCTTTCAGTTATGATGAGATGCTAGTTTACCTTT 617
 DB 579 TCCATAGACCAAGATGATGTTCTGGGCTTTCAGTTATGATGAGATGCTAGTTTACCTTT 638
 QY 618 CTTGACGTTGATTAACCTTTATTTTGGAGAAACGGGCCAGGAAAGATCTATTATCTCGC 677
 DB 639 CTTGACGTTGATTAACCTTTATTTTGGAGAAACGGGCCAGGAAAGATCTATTATCTCGC 698
 QY 678 TATTCACAGGGAACCACTAGGCTTTTATGATTTTCCACCATGCGAGCTGGCTCAG 737
 DB 699 TATTCACAGGGAACCACTAGGCTTTTATGATTTTCCACCATGCGAGCTGGCTCAG 758
 QY 738 AAAATCAAAATGATTTTCTTTAGCACCATAGCAGCTGTTAAGCATGCCAAAAGCCCC 797
 DB 759 AAAATCAAAATGATTTTCTTTAGCACCATAGCAGCTGTTAAGCATGCCAAAAGCCCC 818
 QY 798 GGGACCAAAATTTTGTGCTGCGAGATATGATGATCAAGGGATTTTGGCAAAAAGAA 857
 DB 819 GGGACCAAAATTTTGTGCTGCGAGATATGATGATCAAGGGATTTTGGCAAAAAGAA 878
 QY 858 TTTCTGTATCAGACCAAGATTTCTCAGCAACTGTTTATTTACCTTTGTGGCCAGGTGAT 917
 DB 879 TTTCTGTATCAGACCAAGATTTCTCAGCAACTGTTTATTTTACCTTTGTGGCCAGGTGAT 938
 QY 918 CTTGATCAGATTTGTAGTAATATCATGTTACTTCTGGGTGGATTCACACCAACATATG 977
 DB 939 CTTGATCAGATTTGTAGTAATATCATGTTACTTCTGGGTGGATTCACACCAACATATG 998
 QY 978 AACATGAGCGGAGCAAGTGTATATGCTGCCACACTCTTGTCTGGAACTCTGTGCAAAAT 1037
 DB 999 AACATGAGCGGAGCAAGTGTATATGCTGCCACACTCTTGTCTGGAACTCTGTGCAAAAT 1058
 QY 1038 ATTCTACACTGGAGCCAGCAGTGAATCTGTGAACTCCGGGCAATTTGACTGGGGAGT 1097
 DB 1059 ATTCTACACTGGAGCCAGCAGTGAATCTGTGAACTCCGGGCAATTTGACTGGGGAGT 1118
 QY 1098 GAGACCAAAATCTGGAAATATGCAATCAGCAACTCTCTGTAAGGTACAGAGTCAAGAT 1157
 DB 1119 GAGACCAAAATCTGGAAATATGCAATCAGCAACTCTCTGTAAGGTACAGAGTCAAGAT 1178
 QY 1158 ATGACGGTCCCTACAGCAATGTGGAAGAGTCAAGAGTCAAGCTTTCAAAATCCAGAAAC 1217
 DB 1179 ATGACGGTCCCTACAGCAATGTGGAAGAGTCAAGAGTCAAGCTTTCAAAATCCAGAAAC 1238
 QY 1218 GTGAAATATGCTCTCTGAGGTGAGCCACTCTCTCACTCACTAAGAAATATTCCTGAATGG 1277
 DB 1239 GTGAAATATGCTCTCTGAGGTGAGCCACTCTCTCACTCACTAAGAAATATTCCTGAATGG 1298
 QY 1278 GCTCAGCTGGATTTTCACTCTGGGCTTTGGATGCTCTCACCGTATGTACAAATGAAATCATC 1337
 DB 1299 GCTCAGTGGATTTTCACTCTGGGCTTTGGATGCTCTCACCGTATGTACAAATGAAATCATC 1358
 QY 1338 CATCTGATCAGCAGGAGGAGCAACCTTTCCAGGAGCGGTGTGAGGCCGTATTGTGA 1397
 DB 1359 CATCTGATCAGCAGGAGGAGCAACCTTTCCAGGAGCGGTGTGAGGCCGTATTGTGA 1418
 QY 1398 AGCATCTGACACTGACATCTTAGGACCACTCTCTGAGGATGGGCTAGGACCCATGAA 1457
 DB 1419 AGCATCTGACACTGACATCTTAGGACCACTCTCTGAGGATGGGCTAGGACCCATGAA 1478
 QY 1458 GGCAGAAATTCAGGAGAGCAGACCTAGTATATATTTTTCAGATTCCTCTGCACTTTGGCAC 1517

Db 1479 GGCAGAAATATGAGAGACAGACCTAGTATACATATTTTTCAGATCCCTGCACCTGGCAC 1538
 QY 1518 TAAATCGACACTTACATTTATTTTCTGTAATAATTAAGTACTTATTAGTAAT 1577
 Db 1539 TAAATCGACACTTACATTTATTTTCTGTAATAATTAAGTACTTATTAGTAAT 1598
 QY 1578 AGAGGTTTTGTATGCTATTATATATTTCTACCATCTTGAAGGCTAGTTTACCTGATAGC 1637
 Db 1599 AGAGGTTTTGTATGCTATTATATATTTCTACCATCTTGAAGGCTAGTTTACCTGATAGC 1658
 QY 1638 CAGAAATATCTAGACATTTCTATATCATTTCAAGTAAATCTCTTTAAACACCTATTGT 1697
 Db 1659 CAGAAATATCTAGACATTTCTATATCATTTCAAGTAAATCTCTTTAAACACCTATTGT 1718
 QY 1698 TTTTCTTATAAGCCATATTTTTCAGACACTTAAAGTAAATGCAATTTGGGACAGATAT 1757
 Db 1719 TTTTCTTATAAGCCATATTTTTCAGACACTTAAAGTAAATGCAATTTGGGACAGATAT 1778
 QY 1758 GAGGCTGGAGCTGTGGATTTATTTGACTTTTGAACAAATTAAGCTAGACATTTTCACT 1817
 Db 1779 GAGGCTGGAGCTGTGGATTTATTTGACTTTTGAACAAATTAAGCTAGACATTTTCACT 1838
 QY 1818 TGTTCACACAGACATACACTACCTCCAGGAGCTGAGCTGCTTTAAGGACACACAA 1877
 Db 1839 TGTTCACACAGACATACACTACCTCCAGGAGCTGAGCTGCTTTAAGGACACACAA 1898
 QY 1878 CAAATCAGTGTACAGTATGATGAATCTATTTAAGCATTTCTCAGAAATAGGCCAAG 1937
 Db 1899 CAAATCAGTGTACAGTATGATGAATCTATTTAAGCATTTCTCAGAAATAGGCCAAG 1958
 QY 1938 TTTTATAGTTGCATCTCAGGAGAAATTTTATAGATGTTTATAGATTTCTTCAATAA 1997
 Db 1959 TTTTATAGTTGCATCTCAGGAGAAATTTTATAGATGTTTATAGATTTCTTCAATAA 2018
 QY 1998 TGCATTTCTGCATACAT 2014
 Db 2019 TGCATTTCTGCATACAT 2035

RESULT 4

AAD19220
 ID AAD19220 standard; DNA; 1384 BP.
 XX
 AC AAD19220;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human CG162 (or C59) lipase DNA #1.
 XX
 KW Human; apolipoprotein; lipase; lipoprotein receptor; ALLr; angina;
 KW cardiovascular disease; lipid metabolism; myocardial infarction;
 KW cerebral ischaemia; arterial thrombosis; thrombolytic; antilipemic;
 KW coronary artery thrombosis; cerebral artery thrombosis; stroke;
 KW intracardiac thrombosis; gene therapy; cardiovascular; vasodilator;
 KW neuroprotectant; cerebroprotective; ds.
 XX
 OS Homo sapiens.
 XX
 EN WO200179446-A2.
 XX
 PD 25-OCT-2001.
 XX
 PF 16-APR-2001; 2001WO-US12529.
 XX
 PR 14-APR-2000; 2000US-197137P.
 PR 20-JUN-2000; 2000US-0598042.
 PR 03-AUG-2000; 2000US-0631451.
 PR 22-SEP-2000; 2000US-0667298.
 PR 17-NOV-2000; 2000US-0714936.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Ballinger DG, Loeb D, Montgomery JR, Tang TY, Zhou P, Goodrich R;

PI Liu C, Asundi V, Zhao QA, Wehrman T, Drmanac RT, Ren F, Qian XB;
 PI Wang D;
 XX WPI; 2001-611724/70.
 DR
 XX
 PT Nucleic acids encoding human apolipoproteins, lipases, and lipoprotein
 PT receptor polypeptides, useful for preventing diagnosing and treating
 PT lipid metabolism disorders, thrombosis and cardiovascular diseases -
 XX
 XX Claim 1; Page 160-161; 266pp; English.
 XX
 CC The invention relates to polynucleotides encoding proteins CG122, CG179,
 CC CG95, CG121, CG162, CG27, CG153 and CG168 which are related to proteins
 CC involved in lipid metabolism and cardiovascular disease such as human
 CC apolipoproteins, lipases and lipoprotein receptor proteins. These DNA
 CC and protein sequences are useful for treating or preventing disorders
 CC associated with apolipoproteins, lipases and lipoprotein receptor (ALLr)
 CC expression and for treating lipid metabolism, cardiovascular diseases
 CC and thrombosis. Antibodies against these proteins are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of these sequences. ALLr polypeptides are also
 CC useful for identifying agents (agonists and antagonists) that bind to
 CC them and cells expressing ALLr proteins are useful for identifying a
 CC therapeutic agent for use in treatment of a pathology related to
 CC aberrant expression or physiological interactions of this polypeptide.
 CC Vectors comprising these DNA and protein sequences are also useful for
 CC producing ALLr proteins. The nucleic acids and polypeptides of the
 CC invention are also useful for the treatment of occlusive cardiovascular
 CC diseases, myocardial infarction, cerebral ischaemia, angina, arterial
 CC thrombosis, coronary artery thrombosis and cerebral artery thrombosis
 CC or intracardiac thrombosis and stroke. The nucleotides of the invention
 CC are used in gene therapy. The present sequence is human CG162 (or C59)
 CC Lipase DNA.
 XX
 SQ Sequence 1384 BP; 404 A; 281 C; 328 G; 371 T; 0 other;
 Query Match 65.7%; Score 1343.4; DB 22; Length 1384;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1344; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 18 GGGAAATTCAGCAGGAGAAATATGTGAAGAGTTTAAACCCACAAATTTCTTACTTTA 77
 Db 39 GGGAAATTCAGCAGGAGAAATATGTGAAGAGTTTAAACCCACAAATTTCTTACTTTA 98
 QY 78 GAATTAGTTGTACATTTGGCAGGAGAAATTAATGACAGATGTTGGACATGTTGGAAC 137
 Db 99 GAATTAGTTGTACATTTGGCAGGAGAAATTAATGACAGATGTTGGACATGTTGGAAC 158
 QY 138 TTGTCAACACAGTGGATTTGTCTCACACAGAAATGGAATGTTGGTCTTCTGTTGGCG 197
 Db 159 TTGTCAACACAGTGGATTTGTCTCACACAGAAATGGAATGTTGGTCTTCTGTTGGCG 218
 QY 198 TATATGTTCCAGAGAAATGTGAATTCAGTACATATGCAATTAAGCTGTGGACCCAGAA 257
 Db 219 TATATGTTCCAGAGAAATGTGAATTCAGTACATATGCAATTAAGCTGTGGACCCAGAA 278
 QY 258 GCATTTCATGAATATTAGTGAATCATCAACATCAAGCTATCCCTGTGAGGAATATCAA 317
 Db 279 GCATTTCATGAATATTAGTGAATCATCAACATCAAGCTATCCCTGTGAGGAATATCAA 338
 QY 318 GTCCCAACTGAGATGGGTATATCTTTTCTTTAAACAGGATTCCTCGAGGCTAGTGCAA 377
 Db 339 GTCCCAACTGAGATGGGTATATCTTTTCTTTAAACAGGATTCCTCGAGGCTAGTGCAA 398
 QY 378 CCTAAGAGACAGGTTCCAGGCTGTGGTGTACTTGCAGCATGCGCTAGTTGGAGGTGCT 437
 Db 399 CCTAAGAGACAGGTTCCAGGCTGTGGTGTACTTGCAGCATGCGCTAGTTGGAGGTGCT 458
 QY 438 AGCAACTGGATTTCCAACTGCCCCCAACAAATAGCTGGCTTTCAATCTGGCAGATGCTGGT 497
 Db 459 AGCAACTGGATTTCCAACTGCCCCCAACAAATAGCTGGCTTTCAATCTGGCAGATGCTGGT 518
 QY 498 TTTCAGGTGTGATGGGAAACAGCAGGAGAAACGCTGGTCTCGGAAACACACACACTC 557

Db 519 TTTGACGTGTGGTGGGAAACAGCAGGGGAAACGCTGGTCTCGAAACACAGACACTC 578
 Qy 558 TCATACAGCAAGATGAGTCTGGCTTTCAGTATGATGAGATGCTAGTGTGACCTT 617
 Db 579 TCCATAGCAAGATGAGTCTGGCTTTCAGTATGATGAGATGCTAGTGTGACCTT 638
 Qy 618 CTTGACGTGATAACTTTATTTTGCAGAAACCGGCCAGGAAAGATCTATTATGTCGGC 677
 Db 639 CTTGACGTGATAACTTTATTTTGCAGAAACCGGCCAGGAAAGATCTATTATGTCGGC 698
 Qy 678 TATTACAGGGCAGCAGATGGCTTTATTTGCAATTTTCCACCATGCGAGCTGGCTCAG 737
 Db 699 TATTACAGGGCAGCAGATGGCTTTATTTGCAATTTTCCACCATGCGAGCTGGCTCAG 758
 Qy 738 AAAATCAAATGATTTTGTGTTAGCACCATGACCACTGTTAAGCATGCAAAAAGCCGC 797
 Db 759 AAAATCAAATGATTTTGTGTTAGCACCATGACCACTGTTAAGCATGCAAAAAGCCGC 818
 Qy 798 GGGACCAAATTTTGTGTTGCGCCAGATATGATGATCAAGGGATGTTTGGCAAAAAGAA 857
 Db 819 GGGACCAAATTTTGTGTTGCGCCAGATATGATGATCAAGGGATGTTTGGCAAAAAGAA 878
 Qy 858 TTTCTGATCAGACCACTTCTCAGACAACTTGTATTTTACCTTTGTCGCCAGGTGATT 917
 Db 879 TTTCTGATCAGACCACTTCTCAGACAACTTGTATTTTACCTTTTGGCCAGGTGATT 938
 Qy 918 CTTGATCAGATTTTGTAGTATATCATGTTTCTGGTGGATTCAACACCAATATG 977
 Db 939 CTTGATCAGATTTTGTAGTATATCATGTTTCTGGTGGATTCAACACCAATATG 998
 Qy 978 AACATGAGCCGAGCAAGTATGATGTCGCCACACTTCTGCGAACATCTGTGCAAAAT 1037
 Db 999 AACATGAGCCGAGCAAGTATGATGTCGCCACACTTCTGCGAACATCTGTGCAAAAT 1058
 Qy 1038 ATTCTACACTGGACCCAGGAGTGAATCTGGTGAATCTCGGCAATTTGATCGGGAGT 1097
 Db 1059 ATTCTACACTGGACCCAGGAGTGAATCTGGTGAATCTCGGCAATTTGATCGGGAGT 1118
 Qy 1098 GAGACCAAATCTGGAATAATGCAATCAGCAACTCTCTTAAGGTACAGATCAGAGAT 1157
 Db 1119 GAGACCAAATCTGGAATAATGCAATCAGCAACTCTCTTAAGGTACAGATCAGAGAT 1178
 Qy 1158 ATGACGGTCCCTCAGCAATCTGACAGAGGAGTTCAGACTGGCTTCCAAATCCAGAGAC 1217
 Db 1179 ATGACGGTCCCTCAGCAATCTGACAGAGGAGTTCAGACTGGCTTCCAAATCCAGAGAC 1238
 Qy 1218 GTGAAATGCTGCTCTCTGAGGTGACCAACCTCATCTACCAATGAATATTTCTGAATGG 1277
 Db 1239 GTGAAATGCTGCTCTCTGAGGTGACCAACCTCATCTACCAATGAATATTTCTGAATGG 1298
 Qy 1278 GCTCAGGTGATTTTCATCTGGGTTTGGATGCTCTCAGCGTATGTACATGAATCATC 1337
 Db 1299 GCTCAGGTGATTTTCATCTGGGTTTGGATGCTCTCAGCGTATGTACATGAATCATC 1358
 Qy 1338 CATCTGATGACAGGAGGAGACCA 1362
 Db 1359 CATCTGATGACAGGAGGAGACCA 1383

RESULT 5
 AAF45132
 ID AAF45132 standard; cDNA; 1269 BP.
 XX AC AAF45132;
 XX AC AAF45132;
 DT 30-MAR-2001 (first entry)
 XX Human TANGO 294 ORF.
 DE Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
 KW central nervous system; focal brain disorder; bipolar affective disorder;
 KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;

KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
 KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
 XX neuropsychiatric; psychoactive substance use; anxiety; ss.
 OS Homo sapiens.
 FN WO200077239-A2.
 XX 21-DEC-2000.
 XX 24-MAY-2000; 2000WO-US14858.
 XX 14-JUN-1999; 99US-0333159.
 XX (MILL-) MILLENNIUM PHARM INC.
 XX McCarthy SA, Fraser CC, Sharp JD, Barnes TM;
 XX WPI: 2001-032313/04.
 XX P-PSDB; AAB66065.
 XX TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
 PT screening assays and diagnostic assays and for the treatment of
 PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
 PT disease.
 XX Claim 1; Fig 6; 359pp; English.
 XX The present invention relates to TANGO or INTERCEPT proteins and coding
 CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
 CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
 CC sequences are useful for the treatment of neurological disorders such as
 CC central nervous system (CNS) disorders, CNS-related disorders, focal
 CC brain disorders, global-diffuse cerebral disorders and other
 CC neurological and cerebrovascular disorders. The CNS disorders include
 CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
 CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
 CC autonomic function disorders such as hypertension and sleep disorders,
 CC neuropsychiatric disorders, psychoactive substance use disorders,
 CC anxiety, and bipolar affective disorder.
 XX Sequence 1269 BP; 358 A; 268 C; 309 G; 334 T; 0 other;

Query Match 62.1%; Score 1269; DB 22; Length 1269;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 126 ATGTTGGAAACCTTGTCAAGACAGTGGATTCTCTCACACAGAAATGGAATGTGGCTTCTG 185
 Db 1 ATGTTGGAAACCTTGTCAAGACAGTGGATTCTCTCACACAGAAATGGAATGTGGCTTCTG 60
 Qy 186 ATTCCTGGTGGGTATATCTTCCAGAGAAATGTGAATTCAGTACATATGCCAACTTAAGCT 245
 Db 61 ATTCCTGGTGGGTATATCTTCCAGAGAAATGTGAATTCAGTACATATGCCAACTTAAGCT 120
 Qy 246 GTGGACCCAGAGCAATTCATGAATATTAGTGAATTCATCCAACTCAAGGCTATCCCTGT 305
 Db 121 GTGGACCCAGAGCAATTCATGAATATTAGTGAATTCATCCAACTCAAGGCTATCCCTGT 180
 Qy 306 GAGGAATATGAAGTCCGAACCTGAGATGGGTATATCTCTTTTAAACAGATTCCTCGA 365
 Db 181 GAGGAATATGAAGTCCGAACCTGAGATGGGTATATCTCTTTTAAACAGATTCCTCGA 240
 Qy 366 GGCCTAGTGCACACCTTAAGAGACAGTTCAGAGGCTGTGGTGTACTCTCAGCATGGCCTA 425
 Db 241 GGCCTAGTGCACACCTTAAGAGACAGTTCAGAGGCTGTGGTGTACTCTCAGCATGGCCTA 300
 Qy 426 GTTGGAGGTGCTAGCAACTGGATTTCCAACTGCCCAACAATAGCTGGGTTCATTCTG 485
 Db 301 GTTGGAGGTGCTAGCAACTGGATTTCCAACTGCCCAACAATAGCTGGGTTCATTCTG 360
 Qy 486 GCAGATGCTGGTGTGAGCTGTGGATGGGACAGCAGGGAACGCTGTCTCGAAAA 545

Db 361 GCAGATGCTGTTTTCAGCTGTGGATGGGAACAGACAGGGGAAACGCTGCTCGAAA 420
 Qy 546 CACAAGACACTCTCCATAGACCAAGATGATGTTCTGGCTTTCAGTTATGATGAGATGCT 605
 Db 421 CACAAGACACTCTCCATAGACCAAGATGATGTTCTGGCTTTCAGTTATGATGAGATGCT 480
 Qy 606 AGCTTTGACCTTCTGCTGATGATAAATTTTTCGAGAAACGGGCGCAGGAAAGATC 665
 Db 481 AGCTTTGACCTTCTGCTGATGATAAATTTTTCGAGAAACGGGCGCAGGAAAGATC 540
 Qy 666 TATTATGCTGGCTATTCACAGGACCAACCATGGGCTTTATGCAATTTTCACACATGCCA 725
 Db 541 TATTATGCTGGCTATTCACAGGACCAACCATGGGCTTTATGCAATTTTCACACATGCCA 600
 Qy 726 GAGCTGGCTCAGAAATCAAAATGATTTTTCGACACCCATAGCCACTGCTTAAGAT 785
 Db 601 GAGCTGGCTCAGAAATCAAAATGATTTTTCGACACCCATAGCCACTGCTTAAGAT 660
 Qy 786 GCAAAAGCCCGGACCAAAATTTTTCGCTGCCAGATATGATGATCAAGGATTTT 845
 Db 661 GCAAAAGCCCGGACCAAAATTTTTCGCTGCCAGATATGATGATCAAGGATTTT 720
 Qy 846 GGCAGGATGATTTCTGATCAGATTTTGTAGTAATATCATGTTACTTCTGGGTGATTCAC 965
 Db 781 GGCAGGATGATTTCTGATCAGATTTTGTAGTAATATCATGTTACTTCTGGGTGATTCAC 840
 Qy 966 ACCAACAATATGAATGAGCCAGCAAGTGTATGCTGCCACACACTCTTCTGCTGAACA 1025
 Db 841 ACCAACAATATGAATGAGCCAGCAAGTGTATGCTGCCACACACTCTTCTGCTGAACA 900
 Qy 1026 TCTGTGCAAAATATCTACACTGAGCCAGGCAAGTCAATCTGCTGAACTCCGGGCAATT 1085
 Db 901 TCTGTGCAAAATATCTACACTGAGCCAGGCAAGTCAATCTGCTGAACTCCGGGCAATT 960
 Qy 1086 GACTGGGGAGTCAGACAAAATCTGGAATAATGCAATCAGCCAACTCTCTTAAGGTAC 1145
 Db 961 GACTGGGGAGTCAGACAAAATCTGGAATAATGCAATCAGCCAACTCTCTTAAGGTAC 1020
 Qy 1146 AGAGTCAGATATGAGGTCCTTACAGCAATGTCAGAGGAGTCAGAGCTGGCTTCA 1205
 Db 1021 AGAGTCAGATATGAGGTCCTTACAGCAATGTCAGAGGAGTCAGAGCTGGCTTCA 1080
 Qy 1206 AATCCAGACAGCTGAAAATGCTGCTCTCTGAGGTGACCAACCTCATCTACCATAGAT 1265
 Db 1081 AATCCAGACAGCTGAAAATGCTGCTCTCTGAGGTGACCAACCTCATCTACCATAGAT 1140
 Qy 1266 ATTCCTGAATGGGCTCAGTGATTTTCATCTGGGTTTGGATGCTCTCAGCGTATGATAC 1325
 Db 1141 ATTCCTGAATGGGCTCAGTGATTTTCATCTGGGTTTGGATGCTCTCAGCGTATGATAC 1200
 Qy 1326 AATGAATATCATCTGATGAGCAGGAGGAGACCAACCTTTCCAGGAGCGGTGTAG 1385
 Db 1201 AATGAATATCATCTGATGAGCAGGAGGAGACCAACCTTTCCAGGAGCGGTGTAG 1260
 Qy 1386 GCGGTATG 1394
 Db 1261 GCGGTATG 1269

RESULT 6
 ID ABK86569 standard; cDNA; 1210 BP.
 AC ABK86569;
 XX
 DT 24-SEP-2002 (first entry)
 XX Human cDNA encoding lysosomal acid lipase #1.
 DE
 XX

Human; ss; gene; lysosomal acid lipase; lipid malabsorption illness;
 cystic fibrosis; alcoholism; heart disease; heart attack;
 Wolman disease; cholesterol ester storage disease; brain injury;
 mood disorder; anxiety disorder; thought disorder; volition disorder;
 sleep disorder; neurogenic disorder; myopathic disorder; COPD;
 obesity; cancer; chronic obstructive pulmonary disease; diabetes;
 cardiovascular disorder; Alzheimer's disease; Parkinson's disease;
 anorexia; osteoarthritis; central nervous system disorder;
 peripheral nervous system disorder.
 Homo sapiens.
 Key Location/Qualifiers
 CDS 1..1197
 /*tag= a
 /product= "Lysosomal acid lipase #1"
 WO200236731-A2.
 10-MAY-2002.
 30-OCT-2001; 2001WO-EP12518.
 31-OCT-2000; 2000US-244170P.
 29-MAY-2001; 2001US-293516P.
 (FARB) BAYER AG.
 Xiao Y;
 WPI; 2002-519248/55.
 P-FSDB; AAU99164.
 Novel human lysosomal acid lipase polypeptide, useful for treating
 cancer, diabetes, obesity, chronic obstructive pulmonary disease,
 peripheral or central nervous system disorder or cardiovascular
 disorder
 Claim 1; Fig 1; 126pp; English.
 The invention relates to a purified human lysosomal acid lipase
 polypeptide. Also included are the polynucleotide encoding the
 lipase (or its fragment, derivative, allele or sequence at least 60%
 identical to it), vectors, host cells, a reagent (e.g. an antisense
 oligonucleotide), which binds to the lipase or polynucleotide (used for
 detection and modulating/reducing the lipase activity) and an anti-lipase
 antibody. The lipase and polynucleotide are useful for identifying
 therapeutic agents that either increase or decrease the lipase activity.
 The identified agent, the lipase and polynucleotide are useful for
 treatment of a disease such as lipid malabsorption illness,
 cystic fibrosis, alcoholism, heart disease, heart attack,
 Wolman disease, cholesterol ester storage disease, brain injury,
 mood disorder, anxiety disorder, thought disorder, volition disorder,
 sleep disorder, neurogenic disorder, myopathic disorder, obesity, cancer,
 chronic obstructive pulmonary disease (COPD), diabetes,
 cardiovascular disorder, Alzheimer's disease, Parkinson's disease,
 anorexia, osteoarthritis, a central nervous system disorder and
 a peripheral nervous system disorder. The present sequence is the
 cDNA encoding human lysosomal lipase #1.
 Query Match 55.7%; Score 1138; DB 24; Length 1210;
 Best Local Similarity 100.0%; Pred. No. 1.5e-281;
 Matches 1138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 273 AGTGAATATCCAAATCATCAAGGCTATCCCTGTGAGGAATATGAATGCGCACTGAAGAT 332
 Db 73 AGTGAATATCCAAATCATCAAGGCTATCCCTGTGAGGAATATGAATGCGCACTGAAGAT 132
 Qy 333 GGGTATATCTCTTCTGTTTACAGGATTCCTCGAGGCTAGTCAACCTTAGAAGACAGGT 392
 Db 133 GGGTATATCTCTTCTGTTTACAGGATTCCTCGAGGCTAGTCAACCTTAGAAGACAGGT 192

QY 393 TCCAGGCTGTGGTGTACTGAGCATGGCTAGTGTGAGGTGCTAGCAACTGATTTCC 452
 Db 193 TCCAGGCTGTGGTGTACTGAGCATGGCTAGTGTGAGGTGCTAGCAACTGATTTCC 252
 QY 453 AACCTGCCCCCAAAATAGCTGGGCTTCATCTGGCAGATGCTGGTTTGAAGTGTGATG 512
 Db 253 AACCTGCCCCCAAAATAGCTGGGCTTCATCTGGCAGATGCTGGTTTGAAGTGTGATG 312
 QY 513 GGGAAACAGCAGGGGAAAGCGCTGTCTCGAAACACAAAGACACTCTCCATAGACCAAGAT 572
 Db 313 GGGAAACAGCAGGGGAAAGCGCTGTCTCGAAACACAAAGACACTCTCCATAGACCAAGAT 372
 QY 573 GAGTTCTGGGCTTTCAGTTATGATGAGATGGCTAGGTTTGACCTTCTGCGAGTGAATAAC 632
 Db 373 GAGTTCTGGGCTTTCAGTTATGATGAGATGGCTAGGTTTGACCTTCTGCGAGTGAATAAC 432
 QY 633 TTTATTTTGCAGAAAACGGCCAGGAAAGATCTATTATGTGCGGTATTTACACGGGACCC 692
 Db 433 TTTATTTTGCAGAAAACGGCCAGGAAAGATCTATTATGTGCGGTATTTACACGGGACCC 492
 QY 693 ACCATGGGCTTTATTTGCAATTTTCCACCATGCGCAGAGTGGCTCAGAAAATCAAAATGTAT 752
 Db 493 ACCATGGGCTTTATTTGCAATTTTCCACCATGCGCAGAGTGGCTCAGAAAATCAAAATGTAT 552
 QY 753 TTTGCTTTAGCACCCATAGCCACTCTTAAGCATGCAAAAGCCCGGACCAAAATTTTG 812
 Db 553 TTTGCTTTAGCACCCATAGCCACTCTTAAGCATGCAAAAGCCCGGACCAAAATTTTG 612
 QY 813 TTGCTGCCAGATPATGATGATCAAGGATTTGTTGGCAAAAAGAAATTTCTGTATCAGACC 872
 Db 613 TTGCTGCCAGATPATGATGATCAAGGATTTGTTGGCAAAAAGAAATTTCTGTATCAGACC 672
 QY 873 AGATTCTCAGCAACTTGTATTTACCTTTGTGGCCAGGTGATCTTTGATCAGATTTGT 932
 Db 673 AGATTCTCAGCAACTTGTATTTACCTTTGTGGCCAGGTGATCTTTGATCAGATTTGT 732
 QY 933 AGTAATATCATGTTACTCTGGTGGATTTCAACACCAACAATATGAACATGAGCCGAGCA 992
 Db 733 AGTAATATCATGTTACTCTGGTGGATTTCAACACCAACAATATGAACATGAGCCGAGCA 792
 QY 993 AGTGTATATGCTGCCACACTTCTGCGAACATCTGTGGAACATCTGTGCAAAATATTCTACCTGAGC 1052
 Db 793 AGTGTATATGCTGCCACACTTCTGCGAACATCTGTGGAACATCTGTGCAAAATATTCTACCTGAGC 852
 QY 1053 CAGGCAAGTGAATTTCTGGTGAATCTCGGGCAATTTGACTGGGGAGTGAGACCAAAATCTG 1112
 Db 853 CAGGCAAGTGAATTTCTGGTGAATCTCGGGCAATTTGACTGGGGAGTGAGACCAAAATCTG 912
 QY 1113 GAAATATGCAATCAGCCAACTCTCTTAAGGTACAGAGTCAGAGATATGACGGTCCCTACA 1172
 Db 913 GAAATATGCAATCAGCCAACTCTCTTAAGGTACAGAGTCAGAGATATGACGGTCCCTACA 972
 QY 1173 GCAATGTGACAGGAGGTGAGGCTGCTTTCATAATCCAGAGAGCTGGAATCTGCTC 1232
 Db 973 GCAATGTGACAGGAGGTGAGGCTGCTTTCATAATCCAGAGAGCTGGAATCTGCTC 1032
 QY 1233 TGTAGGTGACCAACTCATCTATACCAATGAATATTTCTGAAATGGGCTCAGCTGGATTTTC 1292
 Db 1033 TGTAGGTGACCAACTCATCTATACCAATGAATATTTCTGAAATGGGCTCAGCTGGATTTTC 1092
 QY 1293 ATCTGGGTTTGGATGCTCTCACCGTATGTACATGAATCAATCCATCTGATCAGCAG 1352
 Db 1093 ATCTGGGTTTGGATGCTCTCACCGTATGTACATGAATCAATCCATCTGATCAGCAG 1152
 QY 1353 GAGGAGACCAACTTTCCAGGGACGGTGTGAGGCGGTATTTGTGAAGCATCTGACACT 1410
 Db 1153 GAGGAGACCAACTTTCCAGGGACGGTGTGAGGCGGTATTTGTGAAGCATCTGACACT 1210

RESULT 7
 ABS55356
 ID ABS55356 standard; cDNA; 1206 BP.

XX ABS55356;
 XX 07-JAN-2003 (first entry)
 XX cDNA encoding human lipase protein.
 DE Human; lipase; cancer; colon; kidney; skin; brain; testis; chromosome 10;
 XX Burkitt's lymphoma; cytostatic; gene therapy; lipase-inhibitor; gene; ss.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 PH 5'UTR 1.13
 FT /*tag= a
 FT CDS 4..1200
 FT /*tag= b
 FT /product= "Human lipase protein"
 FT 3'UTR 1201..1206
 FT /*tag= c
 XX WO200274975-A2.
 XX 26-SBP-2002.
 XX 18-MAR-2002; 2002WO-US08035.
 XX 20-MAR-2001; 2001US-0811825.
 XX (PEKE) PE CORP NY.
 XX Yan C, Di Francesco V, Beasley EM;
 XX WPI; 2002-750560/81.
 XX P-PSDB; ABG71000.
 XX New isolated human lipase peptides and encoding nucleic acids, useful
 PT for diagnosing and treating disorders mediated by human lipase
 PT proteins, e.g. cancer of the colon, kidneys, skin and testis, and
 PT Burkitt's lymphoma -
 XX Claim 4; Fig 1; 79pp; English.
 XX The present invention relates to a new lipase peptide. The methods and
 CC compositions of the present invention are useful for diagnosing and
 CC treating disorders mediated by the human lipase protein, such as cancer
 CC of the colon, kidneys, skin, brain and testis, and Burkitt's lymphoma.
 CC The present nucleic acid sequence represents the human lipase gene
 CC located on chromosome 10. This sequence encodes the human lipase protein
 CC of the invention.
 XX Sequence 1206 BP; 339 A; 260 C; 295 G; 312 T; 0 other;
 SQ
 Query Match 55.1%; Score 1126; DB 24; Length 1206;
 Best Local Similarity 100.0%; Fred. No. 1.8e-278;
 Matches 1126; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 273 AGTGAATATCATCCAAATCAAGGCTATCCCTGTGAGGAATATCAAGTCCGAACCTGAAGAT 332
 Db 76 AGTGAATATCATCCAAATCAAGGCTATCCCTGTGAGGAATATCAAGTCCGAACCTGAAGAT 135
 QY 333 GGGTATATCTTCTGTTAAACAGGATTCCTCAGGCTAGTGCACCTTAAGAAGACAGGT 392
 Db 136 GGGTATATCTTCTGTTAAACAGGATTCCTCAGGCTAGTGCACCTTAAGAAGACAGGT 195
 QY 393 TCCAGGCTGTGGTGTACTGAGCATGGCTAGTGTGAGGTGCTAGCAACTGGATTTC 452
 Db 196 TCCAGGCTGTGGTGTACTGAGCATGGCTAGTGTGAGGTGCTAGCAACTGGATTTC 255
 QY 453 AACCTGCCCAACAATAGCTGGCTTCATTTCTGGCAGATGCTGGTTTGAAGTGTGGATG 512
 Db 256 AACCTGCCCAACAATAGCTGGCTTCATTTCTGGCAGATGCTGGTTTGAAGTGTGGATG 315

XX SQ Sequence 1272 BP; 351 A; 263 C; 311 G; 347 T; 0 other;

Query Match 54.2%; Score 1107.8; DB 24; Length 1272;
Best Local Similarity 97.2%; Pred. No. 8.4e-274;
Matches 1146; Conservative 0; Mismatches 12; Indels 21; Gaps 1;

QY 240 AAGCTGTGGAGCCAGCAAGCATTTCATGAATATAGTGAATCATCAACATCAAGGCTAT 299
DB 94 AATGAAGTCAATCTGAGGTGGATGATATAGTGAATCATCAACATCAAGGCTAT 153
QY 300 CCTGTGAGGATATGAAGTGGCACTGAAGATGGATATATCTTCTGTAAACAGGAT 359
DB 154 CCTGTGAGGATATGAAGTGGCACTGAAGATGGATATATCTTCTGTAAACAGGAT 213
QY 360 CTTGAGGCTAGTCACTAAGACAGAGTTCAGGCTGTGCTGTACTGAGCAT 419
DB 214 CTTGAGGCTAGTCACTAAGACAGAGTTCAGGCTGTGCTGTACTGAGCAT 273
QY 420 GGCCTAGTTGGAGGCTAGCAACTGGATTTCCAACTGSCCAACAATAGCCTGGGCTTC 479
DB 274 GGCCTAGTTGGAGGCTAGCAACTGGATTTCCAACTGSCCAACAATAGCCTGGGCTTC 333
QY 480 ATTCTGGCAGATGCTGTTTTCAGCTGTGGGAGCAGGAGGAAAGCCTGCTCT 539
DB 334 ATTCTGGCAGATGCTGTTTTCAGCTGTGGGAGCAGGAGGAAAGCCTGCTCT 393
QY 540 CGAAACACACAGACACTCTCCATAGACCAAGATGATGTTCTGAGTTATGATGAG 599
DB 394 CGAAACACACAGACACTCTCCATAGACCAAGATGATGTTCTGAGTTATGATGAG 453
QY 600 ATGCTAGTTTACCTTCTGAGTATGAATTTTTCAGAAACAGGCGCAGGAA 659
DB 454 ATGCTAGTTTACCTTCTGAGTATGAATTTTTCAGAAACAGGCGCAGGAA 513
QY 660 AAGATCTATTATGCTGGCTATTTCAGAGGACACCATGGGCTTTATGCTATTTCCACC 719
DB 514 AAGATCTATTATGCTGGCTATTTCAGAGGACACCATGGGCTTTATGCTATTTCCACC 573
QY 720 ATGCCAGCTGGCTCAGAAATCAAAATGATTTTGTCTTAGCACCATAGCCACTGTT 779
DB 574 ATGCCAGCTGGCTCAGAAATCAAAATGATTTTGTCTTAGCACCATAGCCACTGTT 633
QY 780 AAGCATGCAAAAGCCCGGACCAAAATTTTGTCTGCGCAGATATGATCAAGGGA 839
DB 634 AAGCATGCAAAAGCCCGGACCAAAATTTTGTCTGCGCAGATATGATCAAGGGA 693
QY 840 TTGTTTGGCAAAAGAAATTTCTGATCAGACCAAGTTTCTCAGACAACTTTGTTATTAC 899
DB 694 TTGTTTGGCAAAAGAAATTTCTGATCAGACCAAGTTTCTCAGACAACTTTGTTATTAC 753
QY 900 CTTTGTGGCCAGTGTATCTGATCAGATTTGTAGTATATCATGTTACTTCTGGTGA 959
DB 754 CTTTGTGGCCAGTGTATCTGATCAGATTTGTAGTATATCATGTTACTTCTGGTGA 813
QY 960 TTCAACACCAACAATATGAACAT-----GAGCCGAGCAAGTGTA 998
DB 814 TTCAACACCAACAATATGAACATGAATCTATGATGTTTACAGAGCCGAGCAAGTGA 873
QY 999 TATGCTGCCACACTTGTCTGGAACATCTGTGCAAAATATCTACATGAGCCAGGCA 1058
DB 874 TATGCTGCCACACTTGTCTGGAACATCTGTGCAAAATATCTACATGAGCCAGGCA 933
QY 1059 GTGAATTTCTGTGAACTCCGGGCAATTTGACTGGGGAGTGAGCCAAAATCTCGAAAAA 1118
DB 934 GTGAATTTCTGTGAACTCCGGGCAATTTGACTGGGGAGTGAGCCAAAATCTCGAAAAA 993
QY 1119 TGAATTCAGCCAACTCTGTGTAAGGTACAGATGATGAGGTCCTCTACAGCAATG 1178
DB 994 TGAATTCAGCCAACTCTGTGTAAGGTACAGATGATGAGGTCCTCTACAGCAATG 1053
QY 1179 TGGACAGAGGTTCAGGCTTCTCAAAATCCAGAGAGCGTGAATGCTCTCTGAG 1238

DB 1054 TGGACAGAGGTGAGGACTGGCTTTCAATCCAGAGAGCGTGAATGCTGCTCTCTGAG 1113
QY 1239 GTGACCAACCTCATCTACCAATAAGATAATTCCTGAATGGGCTCACGTGATTTTCATCTGG 1298
DB 1114 GTGACCAACCTCATCTACCAATAAGATAATTCCTGAATGGGCTCACGTGATTTTCATCTGG 1173
QY 1299 GGTITGATGCTCTCTCACCGTATGTACAATGAATCATCATCTGATGACGAGGAGGAG 1358
DB 1174 GGTITGATGCTCTCTCACCGTATGTACAATGAATCATCATCTGATGACGAGGAGGAG 1233
QY 1359 ACCAACTTTCCAGGAGCGGTGTGAGCGGCTATTGTGA 1397
DB 1234 ACCAACTTTCCAGGAGCGGTGTGAGCGGCTATTGTGA 1272

RESULT 9
ABS55357
ID ABS55357 standard; DNA; 18554 BP.
XX AC ABS55357;
XX AC
XX AC
DT 07-JAN-2003 (first entry)
XX
DE Genomic DNA encoding human lipase protein.
XX Human; lipase; cancer; colon; kidney; skin; brain; testis; chromosome 10;
KW Burkitt's lymphoma; cytostatic; gene therapy; lipase-inhibitor; gene; ds;
KW single nucleotide polymorphism; SNP.
XX Homo sapiens.
XX

Key Location/Qualifiers
CDS 2001..16554
FT /*tag= a
FT /product= "Human lipase protein"
FT 2001..2072
FT /*tag= b
FT /number= 1
FT 2073..4364
FT /*tag= c
FT /number= 1
FT 4365..4482
FT /*tag= d
FT /number= 2
FT 4483..9149
FT /*tag= e
FT /number= 2
FT 9150..9348
FT /*tag= f
FT /number= 3
FT 9349..10582
FT /*tag= g
FT /number= 3
FT 10583..10692
FT /*tag= h
FT /number= 4
FT 10693..11140
FT /*tag= i
FT /number= 4
FT 11141..11277
FT /*tag= j
FT /number= 5
FT 11278..11409
FT /*tag= k
FT /number= 5
FT 11410..11556
FT /*tag= l
FT /number= 6
FT 11557..12868
FT /*tag= m
FT /number= 6
FT 12869..12940
FT /*tag= n

QY 633 TTTATTTTGCAGAAACGGCCAGGAAAGATCTATTATGCTGGCTATTTCAGAGGCACC 692
 Db |||||
 QY 433 TTTATTTTGCAGAAACGGCCAGGAAAGATCTATTATGCTGGCTATTTCAGAGGCACC 492
 Db |||||
 QY 693 ACCATGGGCTTTATTGCAATTTCCACCATGCCAGAGCTGGCTCAGAAAAATCAAAATGTAT 752
 Db |||||
 QY 493 ACCATGGGCTTTATTGCAATTTCCACCATGCCAGAGCTGGCTCAGAAAAATCAAAATGTAT 552
 Db |||||
 QY 753 TTTGCTTTAGACCCCATAGCCTAGCTTACGATGCAAGGATTTGGCAAAAAGAAATTTCTGTATCAGACC 812
 Db |||||
 QY 553 TTTGCTTTAGACCCCATAGCCTAGCTTACGATGCAAGGATTTGGCAAAAAGAAATTTCTGTATCAGACC 612
 Db |||||
 QY 813 TTTGCTTTAGACCCCATAGCCTAGCTTACGATGCAAGGATTTGGCAAAAAGAAATTTCTGTATCAGACC 872
 Db |||||
 QY 613 TTTGCTTTAGACCCCATAGCCTAGCTTACGATGCAAGGATTTGGCAAAAAGAAATTTCTGTATCAGACC 672
 Db |||||
 QY 873 AGATTTCTCAGACAACTTGATTTTACCTTTGGCCAGGATTTCTGTATCAGATTTGT 932
 Db |||||
 QY 673 AGATTTCTCAGACAACTTGATTTTACCTTTGGCCAGGATTTCTGTATCAGATTTGT 732
 Db |||||
 QY 933 AGTAATATCATGTTTACTTCTGGGTGGATTTCAACCAACCAATATGAACAT----- 982
 Db |||||
 QY 733 AGTAATATCATGTTTACTTCTGGGTGGATTTCAACCAACCAATATGAACATGAATCTCAT 792
 Db |||||
 QY 983 -----GAGCCAGCAAGTGTATATGCTGCCACACTTCTGCGAACATCTGTG 1031
 Db |||||
 QY 793 GGTGTTGTAAGAGCCAGCAAGTGTATATGCTGCCACACTTCTGCGAACATCTGTG 852
 Db |||||
 QY 1032 CAAAATATTCTACACTGGAGCCAG 1055
 Db |||||
 QY 853 CAAAATATTCTACACTGGAGCCAG 876
 Db |||||

RESULT 11.

ABN97383
 ID ABN97383 standard; DNA; 2481 BP.
 XX AC ABN97383;
 XX DT 13-AUG-2002 (first entry)
 XX DE Gene #381 used to diagnose liver cancer.
 XX KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
 XX KW metastatic liver tumor; cytotoxic; expression profile; disease state;
 XX KW disease progression; drug toxicity; drug efficacy; drug metabolism.
 XX OS Homo sapiens.
 XX FN WO200229103-A2.
 XX PD 11-APR-2002.
 XX PF 02-OCT-2001; 2001WO-US30589.
 XX PR 02-OCT-2000; 2000US-237054P.
 XX PA (GENE-) GENE LOGIC INC.
 XX PI Horne D, Alvares C, Perea-Da-Silva S, Vockley JG;
 XX DR WPI; 2002-426119/45.
 XX PT Diagnosing and detecting the progression of liver cancer,
 PT hepatocellular carcinoma or metastatic liver tumor in a patient,
 PT involves detecting the level of expression of two or more genes in a
 PT liver tissue sample -
 XX PS Claim 1; SEQ ID NO 381; 298pp; English.
 CC The invention relates to a novel method for diagnosing and detecting the
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
 CC tumour in a patient, and differentiating metastatic liver cancer from

CC hepatocellular carcinoma in a patient, involving detecting the level of
 CC expression of two or more genes represented in ABN97383-ABN97455 in a
 CC tissue sample. The method of the invention has hepatotropic, and
 CC cytotoxic activity. The method is useful for diagnosing and detecting
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic
 CC liver carcinoma in a patient. The method is useful for identifying
 CC expression profiles which serve as useful diagnostic markers as well as
 CC markers that can be used to monitor disease states, disease progression,
 CC drug toxicity, drug efficacy and drug metabolism.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pat_sequences.

XX Sequence 2481 BP; 692 A; 477 C; 513 G; 799 T; 0 other;

Query Match 25.8%; Score 527.8; DB 24; Length 2481;
 Best Local Similarity 66.6%; Pred. No. 6.2e-125;
 Matches 751; Conservative 0; Mismatches 376; Indels 0; Gaps 0;
 QY 240 AAAGCTGTGGACCCAGAGGATTCATGAATATTAGTGAATATCAATCAAGGCTAT 299
 Db |||||
 QY 112 ACAGCTGTGGATCTGAAACAAACATCATGTGAGTGAATATCTCTTACTGGGATTC 171
 Db |||||
 QY 300 CCCTGTGAGGAATATGAGTGCACCTGAAGATGGGTATATCTTCTGTAAACAGATT 359
 Db |||||
 QY 172 CCTAGTGAAGGATACCTAGTTGAGACAGAGATGATATATTCTGCTTAACCGAATT 231
 Db |||||
 QY 360 CCTGAGGCTTAGTGCAACCTAAGAACAGAGTTCCAGGCTGTGGTGTACTGACGAT 419
 Db |||||
 QY 232 CCTCATGGGAGGAGAAACCATTTCTGCAAAAGGTCCTCAACAGTTGTCTTCTGCAACAT 291
 Db |||||
 QY 420 GGCTGTGAGGATGCTAGCACTGATTTCCAACTGCTCCCAACATAGCTGGGCTTC 479
 Db |||||
 QY 292 GGCTGTGAGGATGCTAGCACTGATTTCCAACTGCTCCCAACATAGCTGGGCTTC 351
 Db |||||
 QY 480 ATTCTGGCAGATGCTGGTGTGATGGGAAACAGAGGAGGAAACGCTGTCT 539
 Db |||||
 QY 352 ATTCTGTGATGCTGGTGTGATGGGAAACAGAGGAGGAAATACCTGTCT 411
 Db |||||
 QY 540 CGAAACACAGACACTCTCCATAGACCAAGATGATTTCTGGGCTTCAGTTATGATGAG 599
 Db |||||
 QY 412 CGAAACACAGACACTCTCCATAGACCAAGATGATTTCTGGGCTTCAGTTATGATGAG 471
 Db |||||
 QY 600 ATGGCTAGTGTGACCTTCTGCAAGTAACTTTATTTTGCAGAAACCGGCCAGAA 659
 Db |||||
 QY 472 ATGGCAAAATATGACCTACCACTTCCATTAATTTCTGAATAAAATCGGCCAAGAA 531
 Db |||||
 QY 660 AAGATCTATTATGCTGGCTATTACAGGGCACCACTGGCTTTATTTGATTTTCCACC 719
 Db |||||
 QY 532 CAAGTGATATTATGCTGGGTCATTTCTCAAGGGCACCACTATAGTTTATAGCATTTTCAG 591
 Db |||||
 QY 720 ATGCCAGAGCTGGCTCAGAAATCAAAATGTATTCTTTAGCACCCTAGCCATGTT 779
 Db |||||
 QY 592 ATCCCTGAGCTGGCTAAAGGATTTAAATGTTTTTGGCTGGCTCTGTGGCTCCGTC 651
 Db |||||
 QY 780 AAGCATGCAAAAGCCCGGAGCAAAATTTTGTGTCGACAGATATGATGATCAGGGA 839
 Db |||||
 QY 652 GCCTTCTGTACTAGCCCTATGGCCAAATAGGACGATTAACAGATATCAATTAAGGAC 711
 Db |||||
 QY 840 TTGTTTGGCAAAAAGAAATTTCTGTATCAGACCAAGATTTCTCAGACAACTTTGTTATTAC 899
 Db |||||
 QY 712 TTATTTGGACAAAGAAATTTCTCCCGAGAGTGGTTTTTGAAGTGGCTGGGTACCCAC 771
 Db |||||
 QY 900 CTTTGTGGCCAGGATTTCTGTATCAGATTTGTAGTAAATATATCATGTTACTTCTGGTGA 959
 Db |||||
 QY 772 GTTTCACACTCATCTACTGAGGAGCTCTGTGAAATCTCTGTTTTTCTTCTGTGGA 831
 Db |||||
 QY 960 TTCAACACCAACAATATGAACATGAGCCGAGCAAGTGTATATGCTGCCACACTCTTGCT 1019
 Db |||||
 QY 832 TTATATGAGAAATTTAAATATGCTAGATGATGATATATCAACACATTTCTCTGCT 891
 Db |||||
 QY 1020 GGAAATCTGTGCAAAATATTCTACATGGAGCCAGGAGTAATTTCTGTGAACTCCGG 1079
 Db |||||

Db 892 GGAACCTCTGTGCAAAACATGTTACATGAGCCAGGCTGTTAAATTCAAAAGTTTCAA 951
 QY 1080 GCATTGACCTGGGGAGTGAGACCAAAATCTGGAAAATGCAATCAGCAACTCCTGTA 1139
 Db 952 GCCTTGACTGGGAGCGTCCAGAAATATTTTCAATCAACAGATTAACCTCCC 1011
 QY 1140 AGGTACAGAGTCAGAGATATGACGCTCCCTACAGCAATGTGACAGGAGTCAAGACTGG 1199
 Db 1012 ACATACATATGTGAAGGACATGCTTGTGCGACTGCGAGTCTGAGCGGGGTCAAGACTGG 1071
 QY 1200 CTTTCAAAATCCAGAGAGCTGAAAATGCTGCTCTCTGAGGTGACCAACTCATCTACCAT 1259
 Db 1072 CTTGAGATGCTTACGAGCTCAATATCTTACTGATCTGATCAGATCAGCACTTGGTGTCCAT 1131
 QY 1260 AAGAAATATTCGAATGGCTCACGTGGATTTTCATCTGGGTTTGGATGCTCTCACCGT 1319
 Db 1132 GAGAGCAATCCGGAATGGGAGCATCTTGACTTCAITGGGGCTGGATGCCCTTGGAGG 1191
 QY 1320 ATGTACAATGAATCATCTCATCTGATGAGGAGGAGGACCAACT 1366
 Db 1192 CTTTATAATAAATTAATTAATCTAATGAGGAATATCAGTGAAGCT 1238

RESULT 12

AAH02906
 ID AAH02906 standard; DNA; 2493 BP.
 AC AAH02906;

15-JUN-2001 (first entry)

Human shear stress-response coding sequence SEQ ID NO: 65.

Human: shear stress-response protein; vascular disease;
 arteriosclerosis; ds.

Homo sapiens.

WO200125427-A1.

12-APR-2001.

02-OCT-2000; 2000WO-JP06840.

01-OCT-1999; 99JP-0280976.

(KYOW) KYOWA HAKKO KOGYO KK.

(NOJI/) NOJIMA H.

Nojima H, Yoshisue H, Obayashi M, Ota T, Kawabata A, Sakurada K;

Kuga T, Sekine S, Nakamura Y, Sugano S;

WPI; 2001-266308/27.

P-PSDB; AAB90783.

DNA sequences, proteins encoded by them and antibodies against them
 useful in diagnosis and treatment of vascular disease caused by

arteriosclerosis -

Claim 20; Page 398-402; 678pp; Japanese.

The present invention provides the protein and coding sequences of a
 number of human shear stress response proteins. These are useful in the
 diagnosis, treatment and screening of vascular diseases caused by
 arteriosclerosis, including heart failure, post-PTCA restenosis and
 hypertension.

Sequence 2493 BP; 692 A; 484 C; 516 G; 801 T; 0 other;

Query Match

Best Local Similarity 25.7%; Score 526.2; DB 22; Length 2493;

Matches 751; Conservative 0; Mismatches 376; Indels 0; Gaps 0;

QY 240 AAAGCTGTGACCCAGAGCAATTCATGAATATTAGTGAATCATCCAAATCAAGGCTAT 299
 Db 119 ACAGCTGTGGATCCTGAACAAACATGAATGTGAGTGAATATATCTTCTTACTGGGATTC 178
 QY 300 CCTGTGAGGAATATGAGTCCCACTGAAGATGGGTATATCTTCTTCTGTTTACAGGATT 359
 Db 179 CCTAGTGGGAAATACCTAGTGGAGACAGAGATGGATATATTCTGTGCTTAAACCGAAT 238
 QY 360 CTTGAGGCTTGTGCAACCTAAGAAAGACAGGTTCCAGGCTGTGGTGTGTTACTGTCAGCAT 419
 Db 239 CCTCATGGGAGGAAGAACCATCTGCAAAAGGTCCTCAACACCATGTTGCTTCTCTGCAACAT 298
 QY 420 GGCTAGTGTGGAGGTGCTAGCACTGATNTCCAACTGCTCCCAAAATAGCTGGGCTTC 479
 Db 299 GGCTGTGTCAGATTTCTAGTAACTGGGTCAAAACCTTGGCAACAGCAGCTGGGCTTC 358
 QY 480 ATTCTGCGCATGCTGTTTGTGAGCTGTGGTGGGAAACAGAGGGGAAACGCTGGTCT 539
 Db 359 ATTCTGCTGATGCTGTTTGTGAGCTGTGGTGGGAAACAGCAGAGGAATACCTGCTCT 418
 QY 540 CGAAACACAAAGACATCTCTCATAGACCAAGATGAGTCTTGGGCTTTCAGTTATGATGAG 599
 Db 419 CGGAAACATAAAGACACTCTCAGTTTCTCAGGATGAATCTGGGCTTTCAGTTATGATGAG 478
 QY 600 ATGCTAGGTTTGCACCTTCTGTCAGTGAATAAATCTTATTTTGCAGAAAACGGGCCAGGAA 659
 Db 479 ATGGCAAAATATGACCTACAGCTTCATTAATCTTCTGAAATAAAATAGCTGGCAGAA 538
 QY 660 AAGATCTATTATGTCGGCTATTACAGGGCACCAATGGGCTTTATTCATTTTCCACC 719
 Db 539 CAAGTGTATTATGTCGGCTATTCTCAAGGCACCACTATAGGTTTATAGCATTTTTCACAG 598
 QY 720 ATGCGAGCTGGCTCAGAAAATCAAAATGTAATTTTGTCTTTAGCACCCATAGCCTGTT 779
 Db 599 ATCCCTGAGCTGGCTAAAAGGATTAATATGTTTTTGGCCCTGGGCTCTGCTGGCTCCGTC 558
 QY 780 AAGCATGCAAAAAGCCCGGACCAAAATTTTGTCTGCCAGATATGATGATCAAGGGA 839
 Db 659 GCCTTCTGTACTAGCTTGGCCAAATTAGGACGATTAACAGATCATCTCTATTAGGAC 718
 QY 840 TTGTTTGGCAAAAAGAAATTTCTGTATCAGACCAATTTCTCAGACAACTTGTATTATAC 899
 Db 719 TTATTTGGAGACAAAGAAATTTCTTCCAGAGTGGCTTTTGAAGTGGTGGGTACCCAC 778
 QY 900 CTTTGTGGCCAGGTGATTTCTTGATCAGATTTGTAGTAATATCATGTTATCTTCTGGGTGA 959
 Db 779 GTTTGCACTCATGTCATCTAGAGGAGCTCTGTGAAATCTCTGTTTTCTTCTGTGTGA 838
 QY 960 TTCAACACCAATATGAAATGAGCCGAGCAGAGTGTATATGCTGCCACACTCTTGTCT 1019
 Db 839 TTTAATGAGAGAAATTTAAATATGTTCTAGAGTGGATGTATATACAAACATCTCTCTGCT 998
 QY 1020 GGAACATCTGTCAAAATATTTCTACACTGGAGCCAGGAGTGAATTTCTGGTGAATCTCCG 1079
 Db 899 GGAATCTGTGCAAAACATGTTTACACTGGAGCCAGCTGTTAAATTTCCAAAGTTTCAA 958
 QY 1080 GCAATTTGACGTGGGAGTGGAGACCAAAATTCGAAAAATGCAATCAGCCAACTCTCTGTA 1139
 Db 959 GCCTTTGACTGGGAAAGCAGTGCAGAAATTTTTCATTACCAACAGAGTTATCTCTCC 1018
 QY 1140 AGGTACAGAGTCAGAGATATGACGCTCCCTACAGCAATGTGACAGAGGTCAGGACTGG 1199
 Db 1019 ACATACATGTGAGGACATGCTTGTGGCAGCTGTCAGTCTGGAGCGGGGTACAGACTGG 1078
 QY 1200 CTTTCAAAATCCAGAGACGTTGAAAATGCTGCTCTCTGAGGTGACCAACCTCATCTACCAT 1259
 Db 1079 CTTGCAATGCTTACGACGTCATATCTTACTGACTCAGATCACCACCTTGGTGTGTCAT 1138
 QY 1260 AAGATATTTCTGAATGGGCTCAGCTGATTTTCTGCTGGGTTTGGATGCTCTCCTCCTCT 1319
 Db 1139 GAGAGCAATCCGGAATGGGAGCATCTTGACTTCAATTTGGGGCTGGATGCCCTTGGAGG 1198
 QY 1320 ATGTACAATGAATCATCTCATCTGATGTCAGCAGGAGGAGACCAACCT 1366

Db 1199 CTTTATATATAAATTATTAATCTAATGAGGAATATCATGTAAGCT 1245

RESULT 13
AAQ42310
ID AAQ42310 standard; DNA; 1378 BP.
XX
AC AAQ42310;
XX
DT 25-MAR-2003 (updated)
DT 20-SEP-1993 (first entry)
XX
DE RGL precursor.
XX
KW Rabbit gastric lipase; RGL; prGLN2.1; fat; bioconversion;
KW hydrolysis; transesterification; ds.
XX
OS Oryctolagus cuniculus.
XX
FH Key Location/Qualifiers
FT misc_RNA 120..1247
FT /*tag= a
FT /note= "claim 1; page 10-11"
XX
PN EP542629-A1.
XX
PD 19-MAY-1993.
XX
PF 12-NOV-1992; 92EP-0403055.
XX
PR 13-NOV-1991; 91FR-0013948.
XX
PA (LJOU) INST RECH JOUVEINAL.
XX
PI Benicourt C, Blanchard C, Junien J;
XX
DR WPI; 1993-161080/20.
DR P-PSDB; AAR37302.
XX
PT Rabbit gastric lipase, its precursor and their DNA - useful for
PT treating conditions linked to gastric lipase deficiency, such as
PT mucoviscidiosis and pancreatic exocrine insufficiency
XX
PS Claim 1; Fig 6; 31pp; French.
XX
CC mRNA was isolated from rabbit gastric mucosal cells, converted to
CC cDNA and fragments cloned in pUC18. Recombinant plasmids were used
CC to transform E.coli MM294 and cells screened for reactivity with
CC probes corresp. to known parts of the RGL sequence. One positive
CC clone contd. plasmid pJ0101 which contd. a 1.35 kb PstI sequence.
CC This fragment was isolated, subjected to PCR and the amplification
CC prod. ligated with the 2.2 kb BglII-SacI fragment of pR276 (contg.
CC a synthetic Tac promoter and transcription stop signal). The
CC recombinant plasmid was designated pRGLN2.1.
CC (Updated on 25-MAR-2003 to correct FN field.)
XX
SQ Sequence 1378 BP; 397 A; 287 C; 276 G; 418 T; 0 other;
Query Match 23.9%; Score 489.2; DB 14; Length 1378;
Best Local Similarity 65.0%; Pred. No. 3.9e-115;
Matches 722; Conservative 0; Mismatches 388; Indels 0; Gaps 0;
OY 250 ACCGAGAGCATTCATGAATATTAATGTAATCCACATCAAGGCTATCCCTGTGAGG 309
Db 136 ACCCTGAAGTGAATATTAATTAATGTAATTCCTACTGGGGATACCCAGTGAAT 195
OY 310 AATATGAGTGCACATCAAGATGGTATATCCCTTCTGTTAAACAGGATTCCTCGAGGCC 369
Db 136 AATATGAGTGTGACTGAGATGGTATATCCCTTCTGTTAAACAGGATTCCTCGAGGCC 255
OY 370 TAGTSCAACCTTAAGAAGACAGGTTCCAGGCTGTGGTGTACTCCAGCATGGCTAGTTG 429

256 AAAAAAATTCAGGAAACAGAGGCCAGAGACCCGCTGTAATTTTTCAGCATGTTGCTTG 315
430 GAGTGTCTAGCAACTGATTTCCAACTGCTCCCAACAAATAGCTGGGCTTCATTCTGGCAG 489
316 CATCAGCCTCAACTGGATCTCCAACTGCTCCCAACAAATAGCTGGGCTTCATTCTGGCAG 375
490 ATGCTGGTTCAGCTGTGATGGGAAACAGAGGGGAAACGCTGCTTCGAAACACA 549
376 ATGCTGGTTCAGCTGTGATGGGAAACAGAGGGGAAATACCTGCTCCAGAGAAACT 435
550 AGACACTCTCATAGACCAAGATGAGTTCTGGGCTTTTCACTATGATGAGATGCTAGGT 609
436 TATACTATTACACAGACTCCGTTGAATCTGGGCTTTTACGTTTGAATGAGTAAAT 495
610 TTGACCTTCCTGCTGATGATTAATTTTTCAGAGAAACGGGCGCAGAGAAAGATCTATT 669
496 ATGACCTTCAGCCACAAATGACTTTCATCTGTAAGGAAACTGGACAGAGAAAGCTCACT 555
670 ATGCTGGCTATTCAGAGGACCAACATGGGCTTTTATGCAATTTTCCACCATGCGCAGC 729
556 ATGTTGGTTCATTCAGGGGACCACTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 615
730 TGGCTTCAGAAATCAAAATGTAATTTTTCAGCACCATAGCCACTGTTAAGCATGCAA 789
616 TGGCTGAAAGAAATCAAAACCTTTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 675
790 AAGCCCGGAGCAAAATTTTGTGCTGCGCAGATATGATGATCAAGGATTTGTTGGCA 849
676 AAGCCTTTGTAACAAACTTAGGTTTATTCCTCCACCATGTTTCAAGATTAATTTGGTG 735
850 AAAAAAATTTCTGTATCAGACACAGATTTCTCAGACAACTTTGTTATTTTACCTTTGTGGCC 909
736 ACMAAATATTCTACCCACAAATTTCTTGTATCAATTTCTTGCCACTCAAGTGTGTTCCC 795
910 AGTGATTTCTGATCAGATTTTGTAGTAATATCATGTTACTTCTGGTGGATTCAGACCA 969
796 GTGAGACACTGAATGTAATTTGCAAGATGCTTATTTATCATTTGTGGCTTTGACAGCG 855
970 ACAATATGAACATGAGCGAGCAAGTGTATATGCTGCCCCACACTCTTGTCTGGAACTCTG 1029
856 CAACCTTGAAATCATGAGTGGCTTGGATGTGACGTGTCACATAATCCGCGCAGGAATCTCAG 915
1030 TCGAAAAATTTCTACATGAGAGCCAGGAGTGAATTTCTGTTGAATCTCGGGCACTTGAAT 1089
916 TTTCAAAACATGCTGCACTGGACCCAGGCTGTAAATCTGGAATTTTCAAGCTTTTAAT 975
1090 GGGGAGTGAAGCAAAATCTGGAATAATGCAATCAGCCAACTCTCTGTAAGGTACAGAG 1149
976 GGGGAGTTCAGCTCAGAAATGATGCACTTCAATCAGCCCACTCCCTACTACATG 1035
1150 TCAGAGATATGAGCTCCCTACAGCAATGTGGAAGAGGCTCAGGACTGGCTTCAAAATC 1209
1036 TGACCGCCATGAATGTCCAATTTGCAAGTGTGGAGTGGTGGCAATGACTGGTGGCTGACC 1095
1210 CAGAGAGCTGAATGCTGCTCTCTGAGGTGACCAACTCATCTACCATGAAGATATTC 1269
1096 CCAAGATGTTGACCTTTTGTCTTCAAAACTTTTCAATCTATTTTACCAAGAGATTC 1155
1270 CTGAATGGGCTCAGCTGATTTTCAATCTGGGTTTGGATGCTCCTCAGCTATGTAATG 1329
1156 TTCCATATAATCATTGATTTTATCTGGCAATGAATGCTCTCTCAAGAGTTTTACAATG 1215
1330 AATCATCTCATGATGAGGAGGAGA 1359
1216 AATATTATTTCTATGATGCAAAAGATAAAA 1245

RESULT 14
AAN60685
ID AAN60685 standard; DNA; 1367 BP.
XX
AC AAN60685;
XX

DT 25-MAR-2003 (updated)
 DT 08-JUN-1991 (first entry)
 XX Sequence encoding pregastric lipase.
 DE Precursor polypeptide; secretion vector; enzyme; ss.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FT sig_peptide 47..103
 FT mat_peptide /*tag= a
 FT /*tag= b
 XX WO8603778-A.
 XX PN 03-JUL-1986.
 XX PF 23-DEC-1985; 85WO-GB00599.
 XX PR 21-DEC-1984; 84GB-0032483.
 XX PR 23-DEC-1985; 85WO-GB00599.
 XX PR 01-JAN-1986; 86GB-0019568.
 XX (BREW-) BREWING RES FOUND.
 PA (TUBB/) TUBB R S.
 PA (CLLT) CELLTech LTD.
 XX Tubb RS;
 XX WPI: 1986-182910/28.
 DR P-PSDB; AAP60724.
 XX New precursor polypeptide of defined sequence - and corresp. DNA
 PT used to transform hosts for prodn. of the polypeptide
 XX Example; Fig 5; 60pp; English.
 XX The inventors claim a DNA sequence encoding AAP60723 linked to a
 CC promoter upstream and a gene for a polypeptide downstream.
 CC Particular examples are the yeast enzyme AMG, the mammalian enzyme,
 CC gastric lipase and the mammalian lymphokine, interferon-alpha2.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX Sequence 1367 BP; 397 A; 284 C; 269 G; 417 T; 0 other;
 SQ
 Query Match 23.0%; Score 470; DB 7; Length 1367;
 Best Local Similarity 64.2%; Pred. No. 3.2e-110;
 Matches 707; Conservative 0; Mismatches 395; Indels 0; Gaps 0;
 QY 244 CTGTGGACCCAGCAGCATTCATGAATATTAGTGAATCATCAACATCAAGGCTATCCCT 303
 DB 123 CTGGAGCCCTGAAGTACTATGAACATTAGTCAGATGATTACTTATTGGGGATACCCAA 182
 QY 304 GTGAGGATATGAATCGCAACTGAAGATGGGTATATCCTTCTCTTTAACAGGATTCCTC 363
 DB 183 ATGAAGAATATGAATGTGACTGAAGATGGTATATCTTGAAGTCAATAGAAATTCCTT 242
 QY 364 GAGGCTAGTGCACCTAAGAGACAGGTTCCAGGCTGTGGTGTACTGACGATGCC 423
 DB 243 ATGGGAAGAAAATTAGGGAATACAGCCAGAGACCTGTGTGTTTTCAGATGGTT 302
 QY 424 TAGTTGGAGGCTAGCACTGGATTTCACCTGCCCAACAAATAGCCCTGGGCTTCATTC 483
 DB 303 TGCTTGCATCAGCCCAAACTGGATTTCACCTGCCGAACAAAGAGCTTGCCTTCATTC 362
 QY 484 TGGCAGATCTGTTTTCAGCTGTGGATGGGGAACAGCAGGGGAACCCCTGGTTCGAA 543
 DB 363 TGGCAGATCTGTTTATGATGTGTGGCTGGGCAACAGCAGAGGAACACCTGGGCCAGAA 422
 QY 544 AACACAAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTCAGTTATGATGAGATGG 603
 DB 423 GAACTTGTACTATTCCACCAGATTGAGTTGAATTTCTGGGCTTTCAGTTTGTGATGAATGG 482
 QY 604 CTAGGTTTGACCTTCTCTGAGTGATAAACTTTATTTTGAGAAAACGGGCCAGGAAAAGA 663
 DB 483 CTAATATATGACCTTCCAGCCCAATCGACTTCATTTGTAAGAAAATCGGACAGAGCAGC 542
 QY 664 TCTATTATGTGGCTATTCCAGGGCACCACCATGGCTTTATTGTCATTTTCCACCATGC 723
 DB 543 TACACTATGTGGCCATTCAGGGCACCACCATGGTTTATTGCTTTTCCACCAATC 602
 QY 724 CAGAGCTGCTCAGAAAATCAAAATGATTTTGTCTTTAGCACCCATAGCACTGTAAAGC 783
 DB 603 CCAGCCTGGCTAAAAGATCAAAAACCTTCTATGCTCTAGCTCTGTGCTGCACTGTGAAGT 662
 QY 784 ATGCAAAAAGCCCGGACCAAAATTTTCTGCTCCAGATATGATGATCAAGGATTTGT 843
 DB 663 ATACAAAAGCCCTTATAACCAAACTTAGATTGTTTCTCAATCCCTCTTCAAGTTTATAT 722
 QY 844 TTGGCAAAAAGAAATTTCTGTATCAGACCCAGAAATTTCTCAGACAACTTGTATTATACCTTT 903
 DB 723 TTGGTGACAAAATATTCTTACCCACACAACTTCTTTGATCAATTTCTTGTACTGAAGTGT 782
 QY 904 GTGGCCAGGTGATTCTTGTATCAGATTGTTAGTAATATATCATGTTACTTCTGGGTGATTCA 963
 DB 783 GCTCCCTGAGATGCTGAATCTCTTTGAGCAATGCTTATTTATATTGTTGGATTG 842
 QY 964 ACACCAACAATATGAACATGAGCCGAGCAAGTGTATATGCTGCCACACACTTGTCTGGAA 1023
 DB 843 ACAGTAAGAACTTTAACACGAGTCTGCTGATGTGTATCTATCATATAATCCAGCAGGAA 902
 QY 1024 CATCTGTGCAAAATATTCTACACTGAGCCAGGCTGAAATTTCTGTGAACTCCGGGCAAT 1083
 DB 903 CTTCTGTTCAAAACATGTTCCATTTGACCCAGGCTGTTAAGTCTGGGAAATTTCCAGCTT 962
 QY 1084 TTGACTGGGGAGTGAACCAAAAATCTGAAAAATGCAATCAGCCAACTCTCTGTAAGGT 1143
 DB 963 ATGACTGGGGAAGCCAGTTCAGAAATAGGATGCACTATGATCAGTCCCAACCTCCCTACT 1022
 QY 1144 ACAGACTCAGATATGACGGTCCCTACAGCAATGTGACAGAGGTGAGGACTGGCTTT 1203
 DB 1023 ACAATGTGACGCCATGAATGTACCAATTTGCAATGTCAGTGTGGAACGGTGGCAAGGACCTGTGG 1082
 QY 1204 CAAATCCAGAAAGACGTGAAAATGCTCTCTGAGTGAACCAACCTCATCTACCAATAAGA 1263
 DB 1083 CTGACCCCCAGATGTTGGCCCTTTGCTTCCAAACTCCCAATCTTATTACCAACAGG 1142
 QY 1264 ATATTCTGAAATGGGCTCAGTGGATTCATCTGGGGTTTGGATGCTCTCACCGTATGT 1323
 DB 1143 AGATTCTCTTTTACAAATCATTGGACTTTATCTGGGCAATGGATGCCCTCAAGAGTTT 1202
 QY 1324 ACAATGAATCATCCATCTGAT 1345
 DB 1203 ACAATGACATTGTTTCTATGAT 1224
 RESULT 15
 AAN60566
 ID AAN60566 standard; DNA; 1367 BP.
 XX AAN60566;
 AC AC
 XX 25-MAR-2003 (updated)
 DT 22-AUG-1991 (first entry)
 XX Sequence encoding human pregastric lipase.
 XX Cystic fibrosis therapy; enzyme; lipase deficiency; ss.
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT CDS 47..103
 FT /*tag= a

PT mat_peptide 104..1243
 FT /*tag= b
 XX W08601532-A.
 XX 13-MAR-1986.
 XX 15-AUG-1985; 85WO-GB00364.
 XX 21-AUG-1984; 84GB-0021210.
 XX 15-AUG-1985; 85WO-GB00364.
 XX 01-JAN-1986; 86GB-000897.
 XX (CLLT) CELLTECH LTD.
 XX (LOWE/) LOWE P A.
 XX Lowe PA;
 XX WPI; 1986-081634/12.
 DR P-PSDB; AAP60658.
 XX
 PT New gastric lipase protein, esp. of human origin - for treating
 PT lipase deficiency, and DNA sequences coding for it
 XX
 PS Disclosure; Fig 3; 39pp; English.
 XX
 CC The inventors claim a pregastric lipase protein and a gene encoding
 CC it. Gastric lipase protein is useful for oral administration to
 CC treat lipase deficiency, e.g. cystic fibrosis or pancreatitis.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX
 XX Sequence 1367 BP; 397 A; 284 C; 269 G; 417 T; 0 other;
 SQ

Query Match 23.0%; Score 470; DB 7; Length 1367;
 Best Local Similarity 64.2%; Pred. No. 3.2e-110;
 Matches 707; Conservative 0; Mismatches 395; Indels 0; Gaps 0;

244 CTGTGGACCCAGAGCATTGATGATATAGTGAATCATCAACATCAAGGTATCCCT 303
 123 CTGGAAGCCCTGAAGTCACTATGACATTAAGTATGATGATTAATTTGGGGATCCAA 182
 304 GTGAGGAATATGAGTGGCACTCAAGATGGGTATATCTCTGTTAAAGGATTCCTC 363
 183 ATGAAGATATGAGTGTGACTGAAGATGTTTATATCTTTGAAGTCAATGATTCCT 242
 364 GAGCCTAGTGCACCTAAGAGACAGGTTCAGGCTGTTGTTACTGCACATGGCC 423
 243 ATGGGAAGAAAATTCAGGGAATACAGGCCAGAGACCTGTTGTTTTCAGCANTGGT 302
 424 TAGTTGAGGTGCTAGCACTGGATTTCCAACTGCCCAACATAGCTGGGCTTCATTC 483
 303 TGCTTGATCAGCCACAACTGGATTTCCAACTGCCCAACATAGCTGGGCTTCATTC 362
 484 TGGCAGATGCTGGTTTGAAGTGGATGGGGAACAGCAGGGGAACGCTGCTCTCGAA 543
 363 TGGCAGATGCTGGTTATGATGTTGGCTGGGCAACAGCAGAGGAAACACCTGGGCCAGAA 422
 544 AACACAAGACACTCTCCATAGCAAGATGATGTTCTGGCTTTTCAGTTATGATGATGG 603
 423 GAACCTGTACTATTCACCAATTCAGTTGATTTCTGGGCTTTTCAGCTTTGATGATGG 482
 604 CTAGTTTGACCTTCCTGTCAGTGAATACTTTATTTTGCAGAAACCGGCCAGGAAAGA 663
 483 CTAATATGACCTTCAGGCCCAATTCGACTTCATTTGTAAGAAACCTGGACAGAGCAGC 542
 664 TCTATTATGCTGCTATTCAGGGCCACCATGGGCTTTATGCAATTTTCCACCATGC 723
 543 TACACTATGTTGGCCATTTCCAGGGCCACCATGGGCTTTATTTGCTTTTCCACCATC 602
 724 CAGAGCTGGCTCAGAAAATCAAAATGATTTTGTCTTTAGCACCCATAGCCACTGTTAAGC 783
 603 CCAGCCTGGCTAAAAGATCAAAACCTTCTATGCTCTAGTCTCTGTTGGCACTGTGAAGT 662

QY 784 ATGCABAAAGCCCGGACCAAAATTTTGTGTCGACAGATATGATGATCAAGGGATTGT 843
 DB 663 ATACAAAAGCCTTATATAACAACTTAGATTGTTCTCAATCCCTCTTCAAGTTTATAT 722
 QY 844 TTGGCAAAAAAGAAATTTCTGTATCAGACCAAGATTCTCAGACAACTTGTATTATTACCTTT 903
 DB 723 TTGGTGACAAAATATTCTACCCACACAACTTCTTGATCAATTTCTTGCTACTGAAAGTGT 782
 QY 904 GTGGCCAGGTGATTTCTTGATCAGATTTGTAGTATATATCATGTTACTTCTGGGTGGATTC 963
 DB 783 GCTCCCGTGAGTCTGAAATCTCCTTTGAGCAATGCTTTATTAATTTGGGATTTG 842
 QY 964 ACACCAACAATATGAACATGAGCGGAGCAAGTGTATGCTGCCACACACTCTTCTGCTGGAA 1023
 DB 843 ACAGTAAGAACTTTAAACACAGAGTCTGTTGATGTATCTATCACATAATCCAGCAGGAA 902
 QY 1024 CATCTGTGCAAAATATTCTACACTGGAGCCAGGAGTGAATTTCTGGTGAACCTCCGGSCAT 1083
 DB 903 CTTCTGTTCAAAACATGTTCCATTGGACCCAGGCTGTTAAGTCTGGGAATTTCAAGCTT 962
 QY 1084 TTGACTGGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCCAACTCCTGTAAAGT 1143
 DB 963 ATGACTGGGGAGGCCAGTTTCAGAAATAGGATGCACTATGATCAGTCCCAACCTCCCTACT 1022
 QY 1144 ACAGAGTCAGAGATATGACGTCCTTACAGCAATGTGGACAGGAGTCAAGACTGGCTTT 1203
 DB 1023 ACAATGTGACAGCCATGAATGTACCAATTTGAGTGTGGAACGGTGGGCAAGACTGTTGG 1082
 QY 1204 CAAATCCAGAAAGACGTGAAAATGCTCTCTGAGGTGACCAACCTCATCTACCATAGA 1263
 DB 1083 CTGACCCCAAGATGTTGGGCTTTTGTCTCCAAAACCTCCCAATCTTATTTTACCAAGG 1142
 QY 1264 ATATCTCTGAATGGGCTCAGTGGATTTCACTCTGGGCTTTGGATGCTCTCCCTCAGCTATGT 1323
 DB 1143 AGATTCTCTTTTACAATCACTTTGGACTTTATCTGGGCAATGGATGCCCCCTCAAGAGTTT 1202
 QY 1324 ACAATGAATCATCCATCTGAT 1345
 DB 1203 ACAATGACATTTTCTATGAT 1224

Search completed: February 19, 2004, 05:05:57
 Job time : 572.905 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 04:48:12 ; Search time 4470.52 Seconds
(without alignments)
11112.438 Million cell updates/sec

Title: US-10-042-431-45

Perfect score: 2044

Sequence: 1 gtgacccacgctccgggg.....aaaaaaagggcgccgc 2044

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

E8T:

1: em_estba.*

2: em_esthum.*

3: em_estin.*

4: em_estmu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_hic.*

9: gb_est1.*

10: gb_est2.*

11: gb_hic.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: em_gsa_hum.*

18: em_gsa_inv.*

19: em_gsa_pln.*

20: em_gsa_vrt.*

21: em_gsa_fun.*

22: em_gsa_mam.*

23: em_gsa_mus.*

24: em_gsa_pro.*

25: em_gsa_rtd.*

26: em_gsa_pbg.*

27: em_gsa_vri.*

28: gb_gsa1.*

29: gb_gsa2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1005.2	49.2	2050	11 AK037091	AK037091 Mus muscu
2	1005.2	49.2	2678	11 AK085719	AK085719 Mus muscu
3	1003.6	49.1	2555	11 AK037214	AK037214 Mus muscu
4	628.2	30.7	2527	11 AK019504	AK019504 Mus muscu

C	5	592	29.0	629	10	BG743024	602632145
C	6	543	26.6	556	13	BX103621	BX103621
	7	505	24.7	954	10	BF135102	BF135102
	8	497.6	24.3	1079	10	BF141567	BF141567
	9	492	24.1	669	10	BF143546	BF143546
	10	475.4	23.3	2138	11	AK088659	AK088659
	11	461.8	22.6	1350	11	AK010093	AK010093
	12	460.2	22.5	1347	11	AK010116	AK010116
	13	460.2	22.5	1349	11	AK009300	AK009300
	14	460.2	22.5	1350	11	AK010139	AK010139
	15	458.6	22.4	1340	11	AK010124	AK010124
	16	458.6	22.4	1343	11	AK009479	AK009479
	17	458.6	22.4	1345	11	AK009413	AK009413
	18	458.6	22.4	1346	11	AK009473	AK009473
	19	458.6	22.4	1346	11	AK009359	AK009359
	20	458.6	22.4	1346	11	AK009523	AK009523
	21	458.6	22.4	1346	11	AK009729	AK009729
	22	458.6	22.4	1348	11	AK010058	AK010058
	23	458.6	22.4	1349	11	AK009474	AK009474
	24	458.6	22.4	1349	11	AK009546	AK009546
	25	458.6	22.4	1349	11	AK009571	AK009571
	26	458.6	22.4	1349	11	AK009573	AK009573
	27	458.6	22.4	1349	11	AK010148	AK010148
	28	458.6	22.4	1350	11	AK009459	AK009459
	29	458.6	22.4	1350	11	AK009525	AK009525
	30	458.6	22.4	1350	11	AK009773	AK009773
	31	458.6	22.4	1350	11	AK010061	AK010061
	32	458.6	22.4	1350	11	AK010125	AK010125
	33	458.6	22.4	1350	11	AK010203	AK010203
	34	458.6	22.4	1351	11	AK009428	AK009428
	35	458.6	22.4	1352	11	AK009431	AK009431
	36	458.6	22.4	1356	11	AK010019	AK010019
	37	458.6	22.4	1356	11	AK010035	AK010035
	38	457.6	22.4	1346	11	AK009437	AK009437
	39	457.6	22.4	1347	11	AK010103	AK010103
	40	457.6	22.4	1349	11	AK010026	AK010026
	41	457	22.4	1346	11	AK009560	AK009560
	42	457	22.4	1347	11	AK010236	AK010236
	43	457	22.4	1348	11	AK010231	AK010231
	44	457	22.4	1349	11	AK009875	AK009875
	45	457	22.4	1350	11	AK009537	AK009537

ALIGNMENTS

RESULT 1
AK037091
LOCUS
DEFINITION
AK037091 2050 bp mRNA linear HTC 05-DEC-2002
Mus musculus adult female vagina cDNA, RIKEN full-length enriched library, clone:993011004 product:BA30415.1 (NOVEL LIPASE) (FRAGMENT) homolog [Homo sapiens], full insert sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

AK037091.1 GI:26331937
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

REFERENCE
AUTHORS

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunari, T., Iashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujitake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, S., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yananaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schram, L. M., Stauble, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bona, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S., and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

5 The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2050)

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Segabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/

FEATURES

Location/Qualifiers

source

1. .2050
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6J"
/db_xref="FANTOM_DB:930111004"
/db_xref="taxon:10090"
/clone="930111004"
/sex="female"
/tissue type="vagina"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
144..1412

CDS

/notes="unnamed protein product; BA30415.1 (NOVEL LIPASE) (FRAGMENT) homolog [Homo sapiens] (SPTR Q96LG2, evidence: FASTY, 88.5%id, 100%length, match=1119)
putative"
/codon_start=1
/protein_id="BAC29699.1"
/db_xref="GI:26331938"
/translations="MSEILSRVTVSHRVEIWLIIIVALLQNVNSGHLPTKAADPE AFNVSEIIKHGYPSEEEVATEDGIIISVNRIPRGQTRLKKEGSPVVLQHLGLG DSNWISLNPNNSLGFIADAGFDVWNGSRNTRKHTLSIDODEWAFSDEMA RFDLPVINFILQTKQKYYVYVSGTGMGFIAPFTPELAHKIMYFALAPIATV KYARSPGTFLILPDMWIKVLFGRQELFTFFRQLFTYLCQMLDQICSNIIILL GGFTNNMNSRANVTVAHTPACTSVONILHWSQVNSGELRAFDMGSETKNOEKQO PTIRYKVRDMVPTAMWTGGQDMLSNPDDVKTLLSEVTLNLIYHKNIPFWAHVDFIWG LDPQRYNEIILHMKQEPNLPQGTCEVL"
2027..2032
/notes="putative"
2050
/notes="putative"

polyA_signal

polyA_site

BASE COUNT 598 a 415 c 461 g 576 t
ORIGIN

Query Match 49.2%; Score 1005.2; DB 11; Length 2050;

Best Local Similarity 84.0%; Pred. No. 1.5e-152;

Matches 1173; Conservative 0; Mismatches 213; Indels 10; Gaps 3;

QY	48	TTTTTAAACCCACAAATCTTCTTACTTTAGAAATTAGTTGTTTACATTGGCAGGAA-AAAA	106
DB	65	TCCTTTCATCTCTGTTCTTCTTCTGACCTAGATGATAGTTGTTACATCCTTAGCAACAAG	124
QY	107	TAATGCGAGATGTTGGACCAATGTTGGAAACCTTGTCAAGACAGTGGATGTTCTCACACAG	166
DB	125	CAAAATGCCGATGTTGGACCAATGTCAGAAATCTTTGTCAAGAGTGTGGACTGTTTGGCAG	184
QY	167	AATGGAATGTCCTTCTGATTTCTGTTGGGTATATGTTCCAGAGAAATGTAATTCAGT	226
DB	185	AGTGGAGATAGGTTCTGATTTCTGTTAGCGTATTTACTCCAGAAATGTAATTCAGT	244
QY	227	ACATATGCCAAATAAAGCTGTGGACCCAGAACGATTCATGAATATTAGTGAATATCCCA	286
DB	245	ACATTTGCCACGAAAGCTGCGGATCCAGAGCAATTCATGAATGTTAGCGAAATCATCAA	304
QY	287	ACATCAAGGCTATCCCTGTGAGGAATATGAAGTCGCAACTGAAGTGGGTATATCCCTTTC	346
DB	305	ACCAAGGTTATCCAGTGAGGATATGAAGTTGCAACGAGATGGGTATATCCCTTTC	364
QY	347	TGTTAAACAGGATTCCTCGAGCCCTAGTGCAACCTTAAGAAGACAGTTTCCAGGCTGTGGT	406
DB	365	TGTGAACAGATCCCTCGGGACAGACACGGTGTAAAGAAGGAGGATCCAGGCCAGTGGT	424
QY	407	GTTACTGCGAGATGCGCTAGTTGGAGGTGTAGCAACTGGATTTCCAACTGCCCAACAA	466
DB	425	GTTACTGCAACATGGTCTTCTTGGGATGCTAGCAACTGGATTTCCAACTGCCCAACAA	484
QY	467	TAGCCTGGGCTTCATTCTTGGCAGATGCTGGTTTTCAGCTGTGGATGGGGAACACAGCGGG	526
DB	485	CAGCCTGGGTTTATTCTTGGCAGATGTCAGGTTTGTATGTGTGGATGGGGAACACAGAGG	544
QY	527	AAACCCCTGGTCTCGAAAACACAGACACTCTCCATAGACCAAGATGATGTTCTGGGCTTT	586
DB	545	AAACCCCTGGTCTCGAAAACCAAGACCCCTCTCCATAGATCAAGATGATGTTCTGGGCTTT	604

Direct Submission		Db	Qy	338
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)				
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.				
Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/ Location/Qualifiers				
1..2678 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="PANTOM.DB:D730026117" /db_xref="taxon:10090" /clone="D730026117" /sex="female" /tissue_type="mammary gland" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stages="10 days lactation, adult" 118..1386 /note="unnamed protein product; BA30415.1 (NOVEL LIPASE) (FRAGMENT) homolog [Homo sapiens] (SPTR 0961G2, evidence: PASTY, 88.5%ID, 100%length, match=1119) putative" /codon_start=1 /protein_id="BAC39517.1" /db_xref="GI:26351761" /translation="MSSEILSRVTVSHRVEIWLILVAYLLQRNVNSGHLPTKAADE AFMVSEIIKHGYPSEYEVATEDGYILSVNIPRGQTRLKKEGSRPVVLLQHLGLG DASWISNLNLSGLFADAGFDVWNGSNGTWSRKHKLTLSDDEFWAFSYDEWA RFDLPVFNILQTKQKVVYVSGQTMGFIAPSTNPELAHKIMYFALAPIATV KYARSPGTFLLPDMILKVLFGROEFLYQTFRQLPFLYLCGQMLDQICSNLILL GGFNTNNMSRANVYVAHTPAGTSVQNLHWSQAVNSGELRAPDWSQETKQEKNG PTPRIYKVRDMVPTAMVTGGQDLNPNDDVDTLLSEVTLNLIYHKNIPWAHVDYFNG LDAPORVYNEIILHMKQEPNLPQGTCTKRVKL" /note="putative" 2678 /note="putative"				
795 a 522 c 555 g 806 t				
polyA_signal				
polyA_site				
795 a 522 c 555 g 806 t				
BASE COUNT				
ORIGIN				
Query Match 49.2%; Score 1005.2; DB 11; Length 2678; Best Local Similarity 84.0%; Pred. No. 1.3e-152; Matches 1173; Conservative 0; Mismatches 213; Indels 10; Gaps 3;				
Qy	48	TTTTTAAACCCCAAAATTTCTTACTTTAGAAATAGTGTGTACATTGGCAGGAA-AAAA	106	
Db	39	TCCTTCAATCTCTTCTTCTTGACTCTAGATGATGTTGTATACATCTTAGCAACAAG	98	
Qy	107	TAAATGCAGATGTGACCATGTGTGAAACCTTGTCAAGACAGTGGATTGTCTCACAG	166	
Db	99	CAAATGCCAGATGTGGACCATCTTGGAATCTTGTCAGAGTGTGGATGTTTGGCAG	158	
Qy	167	AATGGAATGTGGCTTCTGATTCTGGTGGCGTATATGTTTCCAGAGAAATGTGAATCAGT	226	
Db	159	AGTGGAGATATGGCTTCTGATTCTGGTAGCGTATTTACTCCAAAGAAATGTGAATCGGG	218	
Qy	227	ACATATGCCAACTAAGCTGTGGACCCAGAACATTCATGATATAGTGAATCATCCA	286	
Db	219	ACATTTGGCCACGAAGCTGGGATCCAGAGCATTCATGAATGTTAGCGAAATCATCAA	278	
Qy	287	ACATCAAGGCTATCCCTGTGAGGAAATATGAAGTCGCACTGAAGATGGGTATATCTTTC	346	

167 AATGGAAATGCTGCTTCTGATTTCTGTTGGGTATATGTTCCAGAGAAATGTGAATTCAGT 226
 Db
 167 AGTGGAGATATGCTTCTGATTTCTGTTAGCTATTTACTCCAAAGAAATGTGAATTCAGT 226
 227 ACATATGCCAACTAAAGCTGTGACCCAGAGCAATTCATGAATATTAGTGAATATCA 286
 Db
 227 ACATTTGCCCAAGAAAGCTGGATCCAGAGCAATTCATGAATATTAGTGAATATCA 286
 287 ACATCAAGGCTATCCCTGTCAGGAATATGAATGCGCAACTGAAGATGGGTATATCTTTC 346
 Db
 287 ACACAGGTTATCCCAAGTGAAGATGAAGTTGCAACGGAAGATGGGTATCTTTC 346
 347 TGTTACAGGATTCCTCGAGGCTTAGTGCACCTTAAGAGACAGGTTCCAGGCTGTGT 406
 Db
 347 TGTGAACAGAAATCCCTCGGAGACAGACACGTTTAAAGAGGAAGATCCAGGCTGTGT 406
 407 GTTACTCGCATGCTGCTAGTTGAGGTGTAGCACTGCACTGATTTCCAACTGCTCCCAACA 466
 Db
 407 GTTACTCGCAATGCTGCTTCTGGGATGCTAGCACTGATTTCCAACTGCTCCCAACA 466
 467 TAGCTGCGCTTCATTTCTGGCAGATGCTGTTTGTGACGTGTGATGGGAAACAGCAGGG 526
 Db
 467 CAGCTGCGTTTATTTCTGGCAGATGCTGTTTGTGATGTTGATGTTGATGGGAAACAGCAG 526
 527 AAACGCTGCTGCTGAAACACAGACACTCTCCATAGACCAAGATGAGTTCTGGCTTT 586
 Db
 527 AAACACCTGCTGCTGGAAGACCAAGACCTCTCCATAGATCAAGATGAGTTCTGGCTTT 586
 587 CAGTTATGATGAGATGCTGTTGACCTTTCTGCACTGATTAACCTTTATTTTCAGAA 646
 Db
 587 CAGTTATGATGAGATGCTGTTGACCTTTCTGCACTGATTAACCTTTATTTTCAGAA 646
 647 AACGGCCAGAAAGATCTATTATGCTGCTATTCAGGGCACCACTGCTGCTTAT 706
 Db
 647 AACGGCCAGAAAGATGCTATTATGCTGCTATTCAGGGCACCACTGCTGCTTAT 706
 707 TGCAATTTTCCACCTGCGCAGAGCTGCTCAGAAATCAAAATGATTTTGTCTTAGCAC 766
 Db
 707 TGCAATTTTCCACCTGCGCAGAGCTGCTCAGAAATCAAAATGATTTTGTCTTAGCAC 766
 767 CATAGCCTGTTAAGCATGCAAAAGCCCGGACCAAAATTTTGTGCTGCGCAGAT 826
 Db
 767 CATAGCCTGTTAAGCATGCAAAAGCCCGGACCAAAATTTTGTGCTGCGCAGAT 826
 827 GATGATCAAGGATTTGTCGCAAAAGAAATTTCTGATCAGACCACTGATTTCTCAGACA 886
 Db
 827 GATGATCAAGGATTTGTCGCAAAAGAAATTTCTGATCAGACCACTGATTTCTCAGACA 886
 887 ACTGTTATTTACCTTTGTCGCGAGTATCTTGTATCAGATTTGTAGTAAATCATGTT 946
 Db
 887 GCTTTTATTTACCTTTGTCGCGAGTATCTTGTATCAGATTTGTAGTAAATCATGTT 946
 947 ACTTCTGGGTGATTTCAACACCAACAAATATGAATGAGCGGAGCAAGTGTATGCTGC 1006
 Db
 947 ACTTCTGGGTGATTTCAACACCAACAAATATGAATGAGCGGAGCAAGTGTATGCTGC 1006
 1007 CCACACTCTGCTGGAACATCTGTGCAAAATATTCTAATCTGAGCGGAGCAAGTGAATTC 1066
 Db
 1007 CCATACACTCTGCTGGAACATCTGTGCAAAATATTCTAATCTGAGCGGAGCAAGTGAATTC 1066
 1067 TGGTGAATCTGCGGATTTGACTGCGGAGTGAACCAAAATCTGGAATAATGCAATCA 1126
 Db
 1067 TGGTGAATCTGCGGATTTGACTGCGGAGTGAACCAAAATCTGGAATAATGCAATCA 1126
 1127 GCAACTCTGCTGAGTACAGATGATGATGCGTCTCCATCAGCAATGTGACAGG 1186
 Db
 1127 GCAACTCTGCTGAGTACAGATGATGATGCGTCTCCATCAGCAATGTGACAGG 1186
 1187 AGGTGAGTCTGCTTCAATCCAGAGACCTGGAATGCTGCTCTGAGGTGACCA 1246
 Db
 1187 AGGTGAGTCTGCTTCAATCCAGAGACCTGGAATGCTGCTCTGAGGTGACCA 1246

1247 CCTCATCTACCAAGAATATTCTGAATGGGCTCAGTGGAATTTTCATCTGGGTTTGA 1306
 Db
 1247 CCTCATCTACCAAGAATATTCTGAATGGGCTCAGTGGAATTTTCATCTGGGTTTGA 1306
 1307 TGCTCTCACCGTATGATCAATGAATCATCATCTGATGACGAGGAGGAGACCACT 1366
 Db
 1307 TGCTCTCACCGTATGATCAATGAATCATCATCTGATGACGAGGAGGAGACCACT 1366
 1367 TTCCAGGAGCGGTGTGAGGCGGTATTGTGAAGCATCTGACATGAGCATTTAGGACAA 1426
 Db
 1367 TTCCAGGAGCGGTGTGAGGCGGTATTGTGAAGCATCTGACATGAGCATTTAGGACAA 1426
 1427 CCTCTGAGGAGTGGG 1442
 Db
 1427 CCTCTGAGGAGTGGG 1442
 1418 CCCCATGCGGAATGGG 1433
 Db
 1418 CCCCATGCGGAATGGG 1433

RESULT 4

AK019504

LOCUS

DEFINITION

Mus musculus 0 day neonate skin cDNA, RIKEN full-length enriched
 library, clone:4632427C23 product:BA30415.1 (NOVEL LIPASE)
 (FRAGMENT) homolog [Homo sapiens], full insert sequence.

ACCESSION

AK019504.1 GI:12859754

VERSION

HTC; CAP trapper.

KEYWORDS

Mus musculus (house mouse)

SOURCE

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

99279253

10349636

REFERENCE

2

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

20493374

11042159

REFERENCE

3

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,

Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,

Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,

Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,

Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,

Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,

Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)

20530913

11075861

REFERENCE

4

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,

Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,

Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I.,

Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,

Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,

Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,

Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,

Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M.,

Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,

Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,

Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Buit, C.,

Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,

Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,

Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,

Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,


```

SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 629)
AUTHORS    NIH-MGC http://mgi.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: csaphe-remail.nih.gov
            Tissue Procurement: James Cleaver, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
            Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLNL0631 row: i column: 12
            High quality sequence stop: 629.
FEATURES    Location/Qualifiers
            1..629
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:477427"
                /lab_host="DH10B (T1 phage-resistant)"
                /clone_lib="NCI_CGAP Skn3"
                /note="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI;
                Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
                Average insert size 1.5kb. Library constructed by Life
                Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 204 a 118 c 107 g 200 t
ORIGIN
Query Match      29.0%; Score 592; DB 10; Length 629;
Best Local Similarity 100.0%; Pred. No. 9.3e-86;
Matches 592; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1426 ACCTCCTGAGGAGTGGGCTAGGACCCATGAGGAGCAATACGGAGAGAGAGCTAG 1485
DB 629 ACCTCCTGAGGAGTGGGCTAGGACCCATGAGGAGCAATACGGAGAGAGAGCTAG 570
QY 1486 TATACATTTTTCAGATTCCTTGGCACTGGGCACTAAATCCGACACTTACATTTT 1545
DB 569 TATACATTTTTCAGATTCCTTGGCACTGGGCACTAAATCCGACACTTACATTTT 510
QY 1546 TTTCTGTAATTAAGTACTTATTAGTAAATAGAGTTTGTATGCTATATATTCT 1605
DB 509 TTTCTGTAATTAAGTACTTATTAGTAAATAGAGTTTGTATGCTATATATTCT 450
QY 1606 ACCATCTTGAAGGTTAGGTTTACCTGATAGCCAGAAAATATCTAGACATTTCTATATC 1665
DB 449 ACCATCTTGAAGGTTAGGTTTACCTGATAGCCAGAAAATATCTAGACATTTCTATATC 390
QY 1666 ATTCAGGTAATCTCTTTAAACACCTATTGTTTCTATAGCCATATTTTGGAGCA 1725
DB 389 ATTCAGGTAATCTCTTTAAACACCTATTGTTTCTATAGCCATATTTTGGAGCA 330
QY 1726 CTAAGTAAATATGGCAATATGGGACAGATATTAGAGTCTGAGTCTGAGATATTGTTG 1785
DB 329 CTAAGTAAATATGGCAATATGGGACAGATATTAGAGTCTGAGTCTGAGATATTGTTG 270
QY 1786 ACTTTGACAAAATAGCTAGACATTTTACCTGTTGCCACAGAGACATAACACTACCTC 1845
DB 269 ACTTTGACAAAATAGCTAGACATTTTACCTGTTGCCACAGAGACATAACACTACCTC 210
QY 1846 AGGAAGCTGAGCTGCTTTAAGGACACAAACCAACAAATCATGTGTTAGTATGATGAAA 1905
DB 209 AGGAAGCTGAGCTGCTTTAAGGACACAAACCAACAAATCATGTGTTAGTATGATGAAA 150
QY 1906 TCTATGTTAAGCATTTCTCAGAAATAGGCCAAAGTTTATAGTTGATCTCAGGGAAGAAA 1965
DB 149 TCTATGTTAAGCATTTCTCAGAAATAGGCCAAAGTTTATAGTTGATCTCAGGGAAGAAA 90

```

```

QY 1966 TTTTATAGATCTTATGAGTTCCCAATAATGATTCGCTGCTATACATAA 2017
DB 89 TTTTATAGATCTTATGAGTTCCCAATAATGATTCGCTGCTATACATAA 38
RESULT 6
BX103621/c
LOCUS      BX103621 Soares fetal lung NbHL19W Homo sapiens cDNA clone
DEFINITION IMAGP998G21690 ; IMAGE:307244, mRNA sequence.
ACCESSION  BX103621
VERSION     BX103621.1 GI:27845805
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 556)
            Ebert, U., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,
            Radelof, U., Schneider, D. and Korn, B.
            Human UnigeneSet - RZPD3
            Unpublished
            Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
            RZPD; IMAGP998G21690.
            RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
            Human UnigeneSet - RZPD3 (RZPDLIB No.972)
            http://www.rzpd.de/CloneCards/cgi-
            bin/showLib.pl.cgi?response?libNo=972 Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 101
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available royalty-free from RZPD;
            contact RZPD (clone@rzpd.de) for further information. Seq primer:
            M13r, Primer sequence: TTTACACAGGAACAGCATGAC.
FEATURES    Location/Qualifiers
            1..556
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGP998G21690 ; IMAGE:307244"
                /dev_stage="19 weeks"
                /lab_host="DH10B (ampicillin resistant)"
                /clone_lib="Soares fetal lung NbHL19W"
                /note="Organ: lung; Vector: pT73D (Pharmacia) with a
                modified polylinker; Site: 1: Not I; Site 2: Eco RI; 1st
                strand cDNA was primed with a Not I - oligo(dT) primer
                [5'-TGTTACCAATCTGAAGTGGGAGCGGCGCAATTTTTTTTTTTT-3'],
                double-stranded cDNA was size selected, ligated to Eco RI
                adapters (Pharmacia), digested with Not I and cloned into
                the Not I and Eco RI sites of a modified pT73 vector
                (Pharmacia). Library went through one round of
                normalization to a Cot = 5. Library constructed by Bento
                Soares and M. Fatima Bonaldo. This library was constructed
                from the same fetus as the fetal heart library. Soares
                fetal heart NbHL19W."
BASE COUNT 177 a 106 c 93 g 179 t 1 others
ORIGIN
Query Match      26.6%; Score 543; DB 13; Length 556;
Best Local Similarity 99.6%; Pred. No. 7.8e-78;
Matches 554; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1425 AACCTCCTGAGGATGGGGCTAGGACCCATGAGGAGCAATACGGAGAGAGACCTA 1484
DB 556 AACCTCCTGAGGATGGGGCTAGGACCCATGAGGAGCAATACGGAGAGAGACCTA 497
QY 1485 GTATACATTTTTCAGATTCCTTGGCACTTGGCACTAAATCCGACACTTACATTTT 1544

```


Db 496 GTATACATTTTCAGATCCCTGCACCTTGGCCACTAAATCCGACACTTACATTTTACATTTT 437
 QY 1545 TTTTCGTAAATTAAGTACTTATTAGGTAATAGAGGTTTGTGATGCTATTATATATTC 1604
 Db 436 TTTTCGTAAATTAAGTACTTATTAGGTAATAGAGGTTTGTGATGCTATTATATATTC 377
 QY 1605 TACCATCTTGAAGGGTAGGTTTACCTGTATAGCCAGAAATATCTAGACATCTCTATAT 1664
 Db 376 TACCATCTTGAAGGGTAGGTTTACCTGTATAGCCAGAAATATCTAGACATCTCTATAT 317
 QY 1665 CATTCAGGTAATCTCTTTAAACACCTATTCTTTTCTATAGCCATATTTTGGAGC 1724
 Db 316 CATTCAGGTAATCTCTTTAAACACCTATTCTTTTCTATAGCCATATTTTGGAGC 257
 QY 1725 ACTAAGTAAATGGCAATTTGGACAGATATTGAGGCTGAGGCTGTGGATTATGTT 1784
 Db 256 ACTAAGTAAATGGCAATTTGGACAGATATTGAGGCTGAGGCTGTGGATTATGTT 197
 QY 1785 GACTTTGACAAATAGCTAGACATTTTCACTTGTGGCAGAGACATACACTACT 1844
 Db 196 GACTTTGACAAATAGCTAGACATTTTCACTTGTGGCAGAGACATACACTACT 137
 QY 1845 CAGGAAGCTGAGCTGCTTTAAGCAACCAACCAAAATCAGTGTACAGTATGATGAA 1904
 Db 136 CAGGAAGCTGAGCTGCTTTAAGCAACCAACCAAAATCAGTGTACAGTATGATGAA 77
 QY 1905 ATCTATGTAAGCAATCTCAGATAAGGCCAAGTTTATAGTTGAT-CTCAGGGAAGAA 1963
 Db 76 ATCTATGTAAGCAATCTCAGATAAGGCCAAGTTTATAGTTGATGATCTCAGGGAAGAA 17
 QY 1964 AATTTTATAGGATGTT 1979
 Db 16 AATTTTATAGGATGTT 1

RESULT 7
 BF135102
 LOCUS 601780142F1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4008131 5',
 mRNA sequence.
 ACCESSION BF135102
 VERSION BF135102.1 GI:10974142
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 954)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs@mail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LAM9242 row: c column: 12
 High quality sequence stop: 695.
 Location/Qualifiers
 1. .954
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="CZECH II"
 /db_xref="taxon:10090"
 /clone="IMAGE:4008131"
 /tissue_type="tumor, metastatic to mammary"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Lu30"
 /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;

FEATURES source

RESULT 8
 BF141567
 LOCUS 601788654F1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4016413 5',
 mRNA sequence.
 ACCESSION BF141567
 VERSION BF141567.1 GI:10980607
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1079)

Site 2: Sali; transgenic model WNT-1, expression driven by
 MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
 dt. Library constructed by Life Technologies, NIH
 Investigator providing samples: Gilbert Smith, NIH
 BASE COUNT 280 a 222 c 222 g 230 t
 ORIGIN

Query Match 24.7%; Score 505; DB 10; Length 954;
 Best Local Similarity 85.3%; Pred. No. 8.4e-72;
 Matches 598; Conservative 0; Mismatches 100; Indels 3; Gaps 3;
 QY 618 CTGCGAGTGATAAATCTTTATTTGCGAGAAACCGGCGAGAAAGATCTATTATGTCGGC 677
 Db 1 CCAGCTGTGATAAATCTTTATCTTACAGAAACCGGCGAGAAAGATCTATTATGTCGGC 60
 QY 678 TATTACAGGCGCACCCATCGGCTTTATTGCAATTTTCCACCATGCGAGAGCTGGCTCAG 737
 Db 61 TACTCAGAGGCGCACCCATCGGCTTTATTGCAATTTTCCACCATGCGAGAGCTAGTCTAT 120
 QY 738 AAAATCAAAATGATTTTGTGCTGCGAGATATGATCATCAAGGATTTGTTGSCAAAAGAA 797
 Db 121 AAAATCAAAATGATTTTGTGCTGCGAGATATGATCATCAAGGATTTGTTGSCAAAAGAA 180
 QY 798 GGGACCAAAATTTTGTGCTGCGAGATATGATCATCAAGGATTTGTTGSCAAAAGAA 857
 Db 181 GGTACCAAAATTTGCTGCTGCGAGATATGATCATCAAGGATTTGTTGSCAAAAGAA 240
 QY 858 TTTCTGTATCAGACCAAGATTTCTCAGACCAATTTGTTTATTATTTACCTTTGTGSCCAAGGAT 917
 Db 241 TTTTATACCAAGATTTTCTCAGACCAATTTTATTATTTTACCTTTGTGSCCAAGGAT 299
 QY 918 CTTGATCAGATTTGTAGTATATCATGTTTACTTCTGGTGGATTTCAACACCAATATG 977
 Db 300 CTGACCAAAATCTGCGAGCAACATCTTACTCTGGAGGTTTAAACAAACATATG 359
 QY 978 AACATGAGCGGAGCAAGTGTATATGCTGCCCACTCTTCTGCGAACATCTGTGCAAAAT 1037
 Db 360 AACATGAGCGGAGCAAGTGTATATGCTGCCCACTCTTCTGCGAACATCTGTGCAAAAT 419
 QY 1038 ATTCTACACTGGAGCGGAGGAGTGAATCTGGTGAATCTGGGCACTTCACTGGGGAGT 1097
 Db 420 ATTCTCCACTGGAGCGGAGGAGTGAATCTGGGCACTTCACTGGGGAGT 479
 QY 1098 GAGACCAAAATCTGGAATAATGCAATCAGCCAACTCTGTAGGTACAGAGTCAGAGAT 1157
 Db 480 GAGACCAAAATCAGGAGAAATGTAATCAGCCAACTCTTATAAGGTACAAAGTTCAGAT 539
 QY 1158 ATGACGCTCCCTCAGCAATGTGGAAGGAGGTCAGAGTGGCTTTCAAATCCAGAGAC 1217
 Db 540 ATGATGCTCCCAACAGCAATGTGGAAGGAGGTCAGAGTGGCTTTCAAATCCAGAGAT 598
 QY 1218 GTGAATAATGCTCTCTGAGGTCAGCAACCTCATCTACCAAGAAATATTTCTGAATGG 1277
 Db 599 GTGAATAATGCTCTCTGAGGTCAGCAACCTCATCTACCAAGAAATATTTCTGAATGG 658
 QY 1278 GCTCACTGGAATTCATCTGGGCTTTGGATGTCCTCACCG 1318
 Db 659 GGTCACTGGA-ATCATCTGGGCTTTGGATGTCCTCACCG 698

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LHAM9263 row: 1 column: 14
 High quality sequence stop: 663.
 Location/Qualifiers
 1. .1079

FEATURES

source
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="CZECH II"
 /db_xref="taxon:10090"
 /clone="IMAGE:4016413"
 /tissue_type="tumor, metastatic to mammary"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Lu30"
 /note="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; transgenic model WNT-1, expression driven by MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies.
 Investigator providing samples: Gilbert Smith, NIH"
 BASE COUNT 312 a 267 c 263 g 237 t
 ORIGIN
 Query Match 24.3%; Score 497.6; DB 10; Length 1079;
 Best Local Similarity 84.1%; Pred. No. 1.2e-70;
 Matches 621; Conservative 0; Mismatches 109; Indels 8; Gaps 5;
 QY 381 AAGAGACAGGTTCCAGGCTTGTGTACTGACGATGCGCTAGTTGGAGTGTGCTAGC 440
 DB 20 AAGAGGAAGGATCCAGGCCAGTGTGTACTGCAACATGCTCTTCTGGGATGCTAGC 79
 QY 441 AACTGGATTTCCACCTGCCCAACATAGCTGGGCTTCACTTGGCAGATGCTGTTT 500
 DB 80 AACTGGATTTCCACCTGCCCAACATAGCTGGGCTTCACTTGGCAGATGCTGTTT 139
 QY 501 GACGTGTGGATGGGAAACAGCAGGGGAAACGCC-TGGTCTCGAAAAACACAGACACTCTC 559
 DB 140 GATGTGTGGATGGGAAACAGCAGAGGAAACACCGTGTCTCGAAGCACAGACCTCTC 199
 QY 560 CATGACCAAGATGATGTTCTGGGCTTTCAGTTATGATGAGATGCTAGTTTGAACCTTC 619
 DB 200 CATGATCAAGATGATGTTCTGGGCTTTCAGTTATGATGAGATGCTAGTTTGAACCTTC 259
 QY 620 TGCAGTGATAAACTTTATTTTGCAGAAAACGGGCCAGGAAAGATCTATTATGTCGGCTA 679
 DB 260 AGCTGTGATAAACTTTATTTACAGAAAACGGGCCAGGAAAGATCTATTATGTCGGCTA 319
 QY 680 TTCACAGGCAACACCATGGGCTTTATGCAATTTTCCACCATGCGCAGCTGGCTCGAA 739
 DB 320 CTCACAGGGCAACACCATGGGCTTTATGCAATTTTCCACCATGCGCAGCTAGCTCTAA 379
 QY 740 AATCAAAATGATTTTGTGTTAGCACCCATGACCACTGTTAAGCATCAAAAAGCCCGG 799
 DB 380 AATCAAAATGATTTTGTGTTAGCTCTTATGACCACTGTTAATATGCAAGATCTCTG 439
 QY 800 GACCAAAATTTTGTGTTGCGCAGATGATGATCAAGGATGTTTGGCAAAAAGAAAT 859
 DB 440 TACCAAAATTTCTGCTGCTGCCAGATGATGATCAAGGATTTATTTGGCAGACAAGAT 499
 QY 860 TCTGTATCAGACCAATTTCTCAGACACTTTGTTATTTTACCTTTGTGCCAGGATTTCT 919
 DB 500 TTTATACCAGACTAGATTTTTCAGGACGCTTTTATTTTACCTTTGTGGGCAAGATTTCT 559
 QY 920 TGATCAGATTTGTAGTAATATCATGTTACTTCTGGTGGATTCACACCAACCAATATGAA 979

DB 560 CGACCAATCTGCAGCAACATCATCTTACTCTGGGAGGTTTAAACACCAACCATATGAA 619
 QY 980 CATGAGCCGAGCA--AGTGTATATGCTGCCCACTCTTCTGCTGGAACATCTGTG--CAA 1034
 DB 620 CATGAGCAGAGCAACATGCTCTGTGCGCCATAAACCTGCTGGAACGTTCTTGTGCCGA 679
 QY 1035 AATATTCTACACTGGAGCCAGG-CAGTGAATTTCTGGTG-AACTCCGGGCATTTGACTGGG 1092
 DB 680 ACCATTCTCACTGGAGCCAGCGCGTGAATTTCTGGGGCACCTTCTGTCGCTTGTGACTGTG 739
 QY 1093 GGAGTGAGACCAAAATC 1110
 DB 740 GGCAAGCGCAACCAACG 757

RESULT 9

BF143546 569 bp mRNA linear EST 24-OCT-2000
 LOCUS 601790896F1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4021870 5',
 DEFINITION mRNA sequence.
 ACCESSION BF143546
 VERSION BF143546
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 669)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LHAM9277 row: 0 column: 23
 High quality sequence stop: 659.
 Location/Qualifiers
 1. .669

FEATURES

source
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="CZECH II"
 /db_xref="taxon:10090"
 /clone="IMAGE:4021870"
 /tissue_type="tumor, metastatic to mammary"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Lu30"
 /note="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; transgenic model WNT-1, expression driven by MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies.
 Investigator providing samples: Gilbert Smith, NIH"
 BASE COUNT 195 a 140 c 165 g 169 t
 ORIGIN

Query Match 24.1%; Score 492; DB 10; Length 669;
 Best Local Similarity 85.5%; Pred. No. 1.2e-69;
 Matches 571; Conservative 0; Mismatches 95; Indels 2; Gaps 2;
 QY 82 TAGTTGTTACATTTGGCAGGAA-AAAATAAATGCAAGATTTGCAATTTGGAACCTTG 140
 DB 3 TAGTTGTTACATCTCTAGCAACAAAGCAATGCGCATGTTGACCATGTCAGAAATCTTG 62
 QY 141 TCAACAGACTGATTTCTCACAAGATGGAATGTTGGCTTCTGATTTCTGTCGTCAT 200
 DB 63 TCAAGAGTGTGACTGTTTTCGACAGAGTGGAGATATGGCTTCTGATTTCTGTAGCATAT 122

QY 201 ATGTTCCAGAGAAATCTGAATTCAGTACATATGCGCACTAAAGCTGTGACCCAGCA 260
 Db 123 TTAATCCAAAGAAATGTGAATCTGGGACATTTTCCGACAAAGCTGCGGATCCAGAGCA 182
 QY 261 TTCAATCAATATTAGTGAATATCATCAACATCAAGGCTATCCCTGTGAGGAATATGAAGTC 320
 Db 183 TTCAATCAATATTAGTGAATATCATCAACATCAAGGCTATCCCTGTGAGGAATATGAAGTC 242
 QY 321 GCAATTCAGAGATGGTATATCTCTTCTGTAAACAGGATCTCTGAGGCTAGTGAACCT 380
 Db 243 GCAACAGAGAGATGGTATATCTCTTCTGTAAACAGGATCTCTGAGGCTAGTGAACCT 302
 QY 381 AAGAAGACAGATGTTCCAGGCTGTGGTGTACTGACATGAGCTAGTTCGAGGCTAGC 440
 Db 303 AAGAAGACAGATGTTCCAGGCTGTGGTGTACTGACATGAGCTAGTTCGAGGCTAGC 362
 QY 441 AACTGATTTCCAACTCTGCGCAATAGCTGGGCTCATCTCTGCGAGATCTGGTTTT 500
 Db 363 AACTGATTTCCAACTCTGCGCAATAGCTGGGCTCATCTCTGCGAGATCTGGTTTT 421
 QY 501 GAGCTGTGATGGGAAACAGCAGGGAAGCGCTGTCTCGAAACACAAAGACACTCTCC 560
 Db 422 GATGTGTGATGGGAAACAGCAGGGAAGCGCTGTCTCGAAACACAAAGACACTCTCC 481
 QY 561 ATAGACAGAGATGATGTTCTGGGCTTTCAATATGATGATGATGATGATGATGATGAT 620
 Db 482 ATAGATCAAGATGATGTTCTGGGCTTTCAATATGATGATGATGATGATGATGATGAT 541
 QY 621 GCAGTCATAAATTTTATTTTCAGAAACGGGCGGAAAGATCTTATATCTCGGCTAT 680
 Db 542 GCTGTGATTAATTTTATTTTCAGAAACGGGCGGAAAGATCTTATATCTCGGCTAT 601
 QY 681 TCACAGGCGCACCATGCTGCTTTATTTGCAATTTTCCACCATGCGGCTCGAGAA 740
 Db 602 TCACAGGCGCACCATGCTGCTTTATTTGCAATTTTCCACCATGCGGCTCGAGAA 661
 QY 741 ATCAAAAT 748
 Db 662 ATCAAAAT 669
 RESULT 10
 LOCUS AK088659
 DEFINITION Mus musculus 2 days neonate thymic cells cDNA, RIKEN full-length enriched library, clone:E430023A04 product:lysosomal acid lipase 1, full insert sequence.
 ACCESSION AK088659
 VERSION 1 GI:26353707
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE 2
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159
 REFERENCE 3
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,

Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, B., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076861
 4
 Kawai, J., Shingawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staab, J.F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Anoh, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombarts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, J., Kohtsuki, S., and Hayashizaki, Y.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)
 21085660
 11217851
 5
 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6 (bases 1 to 2138)
 Fukuda, S., Furuno, M., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Nomura, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Nomazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, P., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.
 Direct Submission
 Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physiological and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
 COMMENT
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.go.jp/
 URL: http://fantom.gsc.riken.go.jp/
 Location/Qualifiers
 1. 2138
 FEATURES
 source

```

/organism="Mus musculus"
/mol_type="rRNA"
/strain="NOD"
/db_xref="FANTOM DB:E430023A04"
/db_xref="taxon:10090"
/clone="E430023A04"
/cell_type="thymic cells"
/tissue_type="thymus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="2 days neonate"
66..1259
CDS
/note="unnamed protein product; lysosomal acid lipase 1
(MGI:96789, GI:26353708, evidence: BLASTN, 99%,
match=2084)
putative"
/codon_start=1
/protein_id="BAC40484.1"
/db_xref="GI:26353708"
/translation="MQLGLVFFVTIGILSRVPTGVTSVAVDPVNNVTEIMRWGY
PGEHSVLTDGDIYLSIHRIPGRKNHFGKPRVYVYLOHGLLADSNWVNTIDNSL
GFLADAGFVWNGSRNTWSLKHETLSVSQDFWAFSEFMAKYDLPASINYLK
TQEOIYVYVHSGCTIGFIAPSPOMPELAKHIMFLVLPVLSNFGAGPLLOLRUP
DPLLDMFGQKFLPOSAMKLSIHVCTHVMKELCANVFFLLCGNEKNLNMVRD
VYTHCPAGTSVQNMHWGVKVKLQAPDMSEKNEHYNGSPSPYNIKNMELP
TALWSGGRDLADINDITILLITQIPKLVHKNIPWDHLDPIFNGLDAPKLYDEIISL
MKYQ"
BASE COUNT 527 a 485 c 508 g 618 t
ORIGIN

```

```

Query Match 23.3%; Score 475.4; DB 11; Length 2138;
Best Local Similarity 64.1%; Pred. No. 3.6e-67;
Matches 716; Conservative 0; Mismatches 401; Indels 0; Gaps 0;

QY 242 AGCTGTGACGACGAGCAATTCATGAATATTAGTGAATATCCAAATCAAGGCTATCC 301
DB 140 AGCTGTGATCTGAATCAATGAATGTGACGAGATAATCATGCGCTGGGGATATCC 199

QY 302 CTGTGAGGAATATGAATCGCACTGAAGATGGGTATATCTTTCTGTAAACAGATTC 361
DB 200 TGGTAGGNACACTCGGTCTGACAGGAGATGGCTATATTCTGAGTATTCACCAATCCC 259

QY 362 TCGAGGCTPAGTGCACCTTGAAGACAGAGTTCAGGCTGTGGTGTACTGACAGCATGG 421
DB 260 TCGTGGCGGAAGAACCAATTTTGGCAAGGTGCCAGACCAAGTGTGTATCTTCAGCAGG 319

QY 422 CCTAGTTGGAGGTGTAGCACTGATTTCCAACTGCGCCCAACATAGCTGGGCTTCAT 481
DB 320 CTGTGTCGAGATCTTAGTAACTGGGTCAAAACATTGACACAGCAGCTTGGGCTTCCT 379

QY 482 TCTGCGAGATGCTGTTTGTGCTGTGGATGGGGAACAGCAGGGAACGCGCTGTCTCG 541
DB 380 CTTGGCGGATGCTGGCTTTGATGTGTGGATGGAAACAGCAGAGGAACACCTGTCTCT 439

QY 542 ARAACACAGACACTCTCCATAGACAGATGAGTCTTGGCTTTCAGTTATGATGAT 601
DB 440 GAAGCACAAAGACTCTTTCAGTTCTCAGATGAGTCTTGGGCTTCAGTTTGTATGAT 499

QY 602 GGTAGGTTTGACCTTCTGTCAGTGATAACTTTATTTTGCAGAAAACGGGCCAGGAAA 661
DB 500 GGCMAAATATGACCTACCTGCTTCAATTACTACATTTTGAATAAATGCGCCAGACA 559

QY 662 GATCTATTATGTCGGTATTCACAGGCAACCAATGGGCTTTATGCAATTTTCCACCAT 721
DB 560 AATCTATTATGTCGGTCTCTTCAAGGCTGCAACCATAGGTTTCATAGCTTTTTCACAG 619

QY 722 GCCAGAGCTGGCTCAGAAATCAAAATGATTTTCTGCTTAGCACCCTAGCAGCATGTTAA 781
DB 620 GCTGAGTTGGCAAGAGATTTAAATGTTTATGCTTGGCTCCCGTGTGCTCTCAA 679

QY 782 GCATGCAAAAACCCCGGACCAAAATTTTGTGTGTCGCAATATGATGATCAAGGGATT 841
DB 680 TTTTGTCTTCAGGCCCGCTACTTCAATTGGGAGCTTGGCAGACCCCTCTTCTCAAGGAC 739

```

```

QY 842 GTTTGGCAAAAAGAAATTTCTGTATCAGACAGCAATTTCTCAGACAACTTGTATTACT 901
DB 740 GTTTGGACAGAGCAATTTCTCCCTCAGAGTCGATGCTGAGTGGCTGAGCAATTCAGT 799

QY 902 TTGTGGCCAGGTGATTTCTGATCAGATTTGTAGTAGTAAATATCATGTTTCTTGGGGG 961
DB 800 TTGCACCCACGTCATCATGAGGAGCTTTGCGCAACGCTTTCTTCTCTGCTATGTGGCT 859

QY 962 CAACACCAACAATATGAACATGAGCGGAGCAAGTGTATATGCTGCCACACACTTCTGTGG 1021
DB 860 CAATGAGAAGAAATTTAATAATATGCTAGAGTGAATGTATATCAACACACTGCCCTGCGG 919

QY 1022 AACATCTGTGCAAAATTTCTACACTGAGGAGCAGCAGTGAATTTCTGGAATCTCGGGC 1081
DB 920 AACTTCTGTGCAAAACATGTTACACTGGGGTCAAGTCTTTTAAATACCGTAAGCTTCAAGC 979

QY 1082 ATTTGACTGGGGAGTGGAGCAAAATCTGGAATAATGCAATCAGCACTCCGTGTAAG 1141
DB 980 TTTTGTACTGGGAGCAGTGGAAGAAGAAATTTTCTATACACAGAGTTTCTTCCCTC 1039

QY 1142 GTACAGAGTCAGAGATATGACGGTCCCTACAGCAATGTGGACAGGAGTTCAGGACTGGCT 1201
DB 1040 GTACACATATAAAACATGCGTCTACCCACTGCTTTGTGGAGCGGGCGCGGACTGGCT 1099

QY 1202 TTCAATCCAGAGAGCTGAAATGCTGCTCTCTGAGGTGACCACTCATCTACCATAA 1261
DB 1100 GGCAGACATCAATGATCATCATCTTACTGACTCAAAATCCCAAGTTGGTGTATCACA 1159

QY 1262 GAATATTTCTGAATGGGCTCAGTGGATTTTCACTCTGGGTTTGGATGCTCTCACCCTAT 1321
DB 1160 GAATTTCTGAGTGGGACCATCTGGATTTTATCTGGGTTTGGATGCTCCCTCGGAACT 1219

QY 1322 GTACATGAATCACTCATCTGATGACGAGGAGGAG 1358
DB 1220 GTATGATGAATAATCAGTCTAATGAAGAATAACAG 1256

```

RESULT 11 AK010093

LOCUS
DEFINITION

AK010093 1350 bp mRNA linear HTC 05-DEC-2002
Mus musculus adult male tongue cDNA, RIKEN full-length enriched
library, clone:2310067K20 Product:TRIACYLGLYCEROL LIPASE, LINGUAL
PRECURSOR (EC 3.1.1.3) (LINGUAL LIPASE) homolog [Rattus
norvegicus], full insert sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AK010093
GI:12845297
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalisation and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)

2

High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)

3

9279253
PUBMED
10349636

AUTHORS

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

JOURNAL
MEDLINE
PUBMED

9279253
PUBMED
10349636

REFERENCE

1

High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)

2

9279253
PUBMED
10349636

AUTHORS

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

JOURNAL
MEDLINE
PUBMED

9279253
PUBMED
10349636

REFERENCE

1

High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)

2

9279253
PUBMED
10349636

source	1. .1347	/organism="Mus musculus"	Qy	790	AAAGCCCGGAGCAAAATTTTCTGCTGCCAGATATGATGATCAAGGATGCTTTGGCA	849
	/mol_type="mRNA"		Db	670	AAAGTCCCTTTAAAGAGATTTCACTTATTCCTAAGTTCTTCTCAAGGTCATATTTGGTA	729
	/strain="C57BL/6J"		Qy	850	AAAAAGAAATTTCTGTATCAGACAGAGATTTCTCAGACAACTTTGTATTTACCTTTGTGGCC	909
	/db_xref="FANTOM DB:2310068005"		Db	730	ACAAAATGTTTCATGCCCCCAACTTACTTAGATCAAATTTCTTGATCGGAAGTGTGCTCAC	789
	/db_xref="MGI:1895801"		Qy	910	AGGTGATCTTGTATCAGATTTGTAGTAATATCATGTACTTCTGGGTGATTCACACCA	969
	/db_xref="taxon:10090"		Db	790	GGAGCTGTAGATCTTCTCTGAGCAACGCTTTATTCATCTTCTGTGATTTGACAAGA	849
	/clone="2310068005"		Qy	970	ACAAATGAACATGAGCCGAGCAAGTGTATATGCTGCCACACTCTTGTGTGAACATCTG	1029
	/sex="male"		Db	850	AAAAATTTAAATGTGAGTGGCTTTGGTGTGTATCTAGGGCAATATCCAGCAGGAACATCTA	909
	/tissue_type="tongue"		Qy	1030	TGCAAAATATTTACACTGAGCCGAGCAGTGAATTTCTGGTGAATCTCCGGGCAATTTGACT	1089
	/clone_lib="RIKEN full-length enriched mouse cDNA library"		Db	910	CTCAAGACCTTTTCCACTGGGCACAGCTTGTCTAAATCTGGGAAGCTTCAAGCCTATAACT	969
	/dev_stage="adult"		Qy	1090	GGGGGAGTGAAGCAAAATTTCTGAAAAATGCAATCAGCCAACTCTCTGTAAGGTACAGAG	1149
	51. .1238		Db	970	GGGGAAGTCCATTACAGAACATGTTACACTACAAAGAGCCCTCCCTACTATGATG	1029
CDS	/note="unnamed protein product; TRIACYLGLYCEROL LIPASE, LINGUAL PRECURSOR (EC 3.1.1.3) (LINGUAL LIPASE) homolog [Rattus norvegicus] (SWISSPROT P04634, evidence: FASTV, 88.4%ID, 100%length, match=1184) putative"		Qy	1150	TCAGAGATATGAGCGTCCCTACAGCAATGTGGACAGAGTCAAGTCTGGCTTTCAATC	1209
	/protein_id="BAB26711.1"		Db	1030	TGTCAAGCATGACCGTGCCTAAATTTGCAAGTGTGGAAAGTGGCCATGATCATCTGGCTGATC	1089
	/db_xref="GI:12845335"		Qy	1210	CAGAGAGCTGAAATGCTGCTCTCTGAGTGAACCAACTCATCTACCAATAGAAATATTC	1269
	/db_xref="MGI:1914967"		Db	1090	CCCAAGATGTCGCAATGCTCTCCCAACTCCCAACCTTCTGTACCATAGGATTC	1149
	/translation="MWLLVTSVLSAFGAGHGLFGKLGKPKNPERNMVNSOMITWYGVPSSEYVWEDGVILGVIRPYGKNSENIGKRPVAYLOHGLIASNTWITLNPNSLAFLADAGVDVWLGNSRGNTWSEKNVYSPDSVEFWAFSDEWAKDLPATIDFVQKTGQEKHYHVGSGTIGFIAPSTNPALAKIKRFYALAPVATVITTESPFKKISLIPKFLKVIKGNMFPNHYLDQFLGTEVSRKLLDLNLFIFCGFDKKNLNVRFVYLNHPAGTSTODLFWAQLAKSGKQAYNWSPLQNNLHYNKTPPYDYSAVTVPVAVWNGDILADPQDVAVMLPKPLNLYHKEILPYNHLDFIAMDAPQEVVNEIVTMABD"		Qy	1270	CTGAATGGGCTCAGCTGGATTTTCATCTGGGTTTGGATGCTCTCCCTGATGTACATG	1329
BASE COUNT	393 a 290 c 284 g 380 t		Db	1150	TTCCCTACATACCTGAGCTTTCATCTGGCGATGGATGGCTCAAGAGGTTTACAATG	1209
ORIGIN			Qy	1330	AAATCATCCATCTGATG 1346	
			Db	1210	AGATAGTTACCATGATG 1226	
Query Match	22.5%; Score 460.2; DB 11; Length 1347;					
Best Local Similarity	63.7%; Pred. No. 1.2e-64;					
Matches	699; Conservative 0; Mismatches 398; Indels 0; Gaps 0;					
Qy	250	ACCAGAAGCATTCATGAATATTAGTGAATCATCAACATCAAGGCTATCCCTGTAGG	309			
Db	130	ACCTGAGCAACATGATGTTAGTCAGATGATACTTACTGGGGATTCAGAGTAGG	189			
Qy	310	ATATGAAGTCGCACTGAAGATGGGTATATCTCTTCTGTAAACAGGATTCCTCGAGCC	369			
Db	190	AAATGAAGTTGTTACTGAAGATGGCTACATCTGGGGGTCTATAGAATTCCTTATGGGA	249			
Qy	370	TAGTGAACCTAAGAAGCAGGTTCCAGGCTGTGTGTACTGAGCAGTGGCTAGTTG	429			
Db	250	AGAAAATTCAGAAATATCGCAAGAGACCTGTGGCATTTTTCAGCATGTTTGAATG	309			
Qy	430	GAGGTGCTAGCACTGGATTTCCAACTGCCCAACAATAGCTGGGCTTCATTTCTGGCAG	489			
Db	310	CATCAGCCCAAACTGGATTACAAATCTGCCAAACAACAGCTGGCCCTTCTATTCAGCAG	369			
Qy	490	ATGCTGTTTTCAGCTGTGGTGGGAGACAGCAGGGAACGCTGTCTCGAAACACA	549			
Db	370	ATGCTGGCTATGATGTGGCTGGGGAACAGTCGAGGGAATACATGGTCCCGGAAAAATG	429			
Qy	550	AGACACTCTCCATAGACCAAGATGATTTCTGGGCTTTTCAGTTATGATGAGATGGCTAGGT	609			
Db	430	TATACTATTACCAAGACTCAGTTGAATTTCTGGGCTTTTCAGCTTTTGTATGAATGGCTAAAT	489			
Qy	610	TTGACCTTCTCAGTGATAAACTTTATTTTCAGAAAACGGCCAGGAAGATCTATT	669			
Db	490	ATGACCTTCAGGCCACCATAGACTTTCATTTGTACAGAAAACCTGGAACAGAGATACACT	549			
Qy	670	ATGCTGGCTATTACAGGGCCACCACCATGGGCTTTTATGTCATTTTCCACCATGCCAGAGC	729			
Db	550	ATGTTGGTCACTCAGGCCACCACCTATCGTTTATGGCTTTTCTACCAATCTGCTC	609			
Qy	730	TGCTCAGAAATCAAAATGATTTTGTCTTTAGCACCACCATGCCACTGTTAAGCATGCAA	789			
Db	610	TGCTAAAAAATCAAGAGGTTTATGCAATAGTCCAGTTCTGCTACTGTGAAGTATACAG	669			
AK009300	1349 bp	mRNA	linear	HTC 05-DEC-2002		
Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310011G18 product:TRIACYLGLYCEROL LIPASE, LINGUAL PRECURSOR (EC 3.1.1.3) (LINGUAL LIPASE) homolog [Rattus norvegicus], full insert sequence.						
AK009300	GI:12844011					
HTC; CAP trapper.						
Mus musculus (house mouse)						
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.						
Carninci, P. and Hayashizaki, Y.						
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.						
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes						
Genome Res. 10 (10), 1617-1630 (2000)						
20493374						
11042159						
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,						

GAGAGAGAGATTCTCGAGTTAATAATTAATCCCCCCCCCCC 3']. cdna was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.

FEATURES

Location/Qualifiers
1. .1350

```
1. 1350
  /organism="Mus musculus"
  /mol_type="mRNA"
  /db_xref="C57BL/6J"
  /db_xref="FANTOM DB:2310069P19"
  /db_xref="PMI:1998377"
  /db_xref="taxon:10090"
  /clone="2310069P19"
  /sex="male"
  /tissue_type="tongue"
  /clone_lib="RIZKEN full-length e"
  /dev_stage="adult"
```

```

/ note="unnamed protein product; TRIACYLGLYCEROL LIPASE,
LINGUAL PRECURSOR (EC 3.1.1.3) (LINGUAL LIPASE) homolog
[Rattus norvegicus] (SWISSPROT P04634, evidence: FASTA,
88.4%ID, 100%length, match=1184)
53..1240
/dev_stage="adult".

```

```

putative"
/codon_start=1
/protein_id="B2626725.1"
/db_xref="GI:12845372"
/db_xref="MG1:1914367"
/translation="MMLLVTVSLGAFGAGHGLFQKLGPKNPNEANMVNQMIYTWGVY
SESYEVTDGTYLGVYRIPYQKNSNIKKRPVAYLQHGILIASATNITNLPNNSLPA
FLLADAGYDWLGNSSKSNKYVSPDSVEFWAFSPDEMAKYDIPATIDFVQTK
GQKHIVHGVSQQTGTGTFATFNSPALAKI KR FYALPAKYTESPFKSLIPK
FLIKVYFGKMFMEQDQGLTGVCSRELLIDLCNALNITFCGDFKQNNISRVFPI
VLGNHPAGTSTODLFHWAQLQSGKGLQANWMSPLQNMLNQKTPPYDVSAMTPI
AVNNGGHDILADPDQDVAMLLPKLPLNLLYKHEILPYNHLDFEWANDA PQEYVNEIVTMMI
ARD"

```

Qy	730	TGCGTCAGAAAATCAAAATGTA	TTTTGCTTTTACGACCCCATAGCCACTGTTTAAAGCATGCAA	789
Db	612	TGCGCTAAAAAAATCAAGAGGTTTTATGCAATTAGCTCCAAATTGCTACTGTGAAGTATACAG		671
Qy	790	AAAGCCCGGGGACCAAAATTTTGTGTGTCGCAGATATGATGATCAAGGGGATCTTTGGCA		849
Db	672	AAAGTCCCTTTTAAAAAGATTTCACTTATTCTTAAAGTTTCTTCTCAAGGTGATATTGGTA		731
Qy	850	AAAAGNAATTTCTGTATCAGACAGATTTCTCAGACAACTTGTATTTTAACTTTGTGGCC		909
Db	732	ACAAAATGTTTCATGCCCCCAACTACTTAGATCAATTTCTTGTGTAGCGAAGTGTGCTCAC		791
Qy	910	AGGTGATTTCTTGATCAGATTTGTAGTAATAATCATGTTACTTCTGGGTGGATTCAAACCA		969
Db	792	GGGAGCTGCTAGATCTTCTCTGCAGCAACCGCTTATTCACTTCTGTGGATTTGACAAAGA		851
Qy	970	ACAATATGAACATGAGCCGAGCAAGTGATATGCTGCCACACTCTTGTCTGGGAACATCTG		1029
Db	852	AAAACCTTAATGTAGTGCCTTTGTATGTGTATCTAGGGCATATTCAGCAGGAACATCTA		911
Qy	1030	TGCAAAATAATTCACACTGGAGCCAGCAGTGAATTTCTGGTGAACCTCCGGGCAATTTGACT		1089
Db	912	CTCAAGACCTTTTCCACTGGGCACAGCTTGCTTAAATCTGGGAAGCTTCAAGCCTATAACT		971
Qy	1090	GGGGAGTGAGACCAAAAATCTCGAAAAATGCAATCAGCCAACTCTCTGTAAGGTACAGAG		1149
Db	972	GGGGAGTCCATTTACAGAACATGTTACATACATCAGAAAACGCTCCCTACTATGATG		1031
Qy	1150	TCAGAGATATGACGGTCCCTACAGCAATGTGGACAGGAGTCAAGACTGGCTTTCAAATC		1209
Db	1032	TGTGAGCCATGACCGGTGCCAAATTTGCAGTGTGGAAACGGTGGCCCATGACATCCTCGCTGATC		1091
Qy	1210	CAGAAGACGTGAAAATGCTGCTCTCTGAGTGCACCAACCTCATCTACCAATAAGAATATTC		1269
Db	1092	CCCAAGATGTGCAATGCTGCTTTCCCAAACTCCCAACCTTTTGTACCAATAAGGAGATTC		1151
Qy	1270	CTGNAATGGGCTCAGTGGATTTTCATCTGGGGTTTGGATGCTCTCTCACCGTATGTACAAATG		1329
Db	1152	TTCCCTACAATACCTTGGACTTCATCTGGCGGATGGATGGCCCTCAAGAGGTTTACAATG		1211
Qy	1330	AAATCATCCATCTGATG		1346
Db	1212	AGATAGTTACCATGATG		1228

RESULT 15
AK010124
LOCUS
DEFINITION

AK010124 Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310069D09 product:TRACYGLYCEROL LIPASE, LINGUAL PRECURSOR (EC 3.1.1.3) (LINGUAL LIPASE) homolog [Rattus norvegicus], full insert sequence

ACCESSION AK010124
VERSION AK010124.1 GI:12845346
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

MUS MUSCULUS
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
PUBMED

Db 546 ATGTTGGTCACTCTCAGGACCACTATCGTTTATTGGCTTTTCTACCAATCCTGCTC 605
 Qy 730 TGGCTCAGAAAATCAAAATGTATTTTCTTTAGCACCATAGCCACTGTTTAAGCATGCAA 789
 Db 606 TGGCTAAAAAATCAAGAGSTTTTATGTCATTAGTCCAGTTGCTACTGTGAAGTATACAG 665
 Qy 790 AAAGCCCGGAGCAAAATTTTGTGCTGCCAGATGATGATCAAGGGATTTTGGCA 849
 Db 666 AAGTCCCTTTAAAAGANTTCACTTATCTTAAGTTTCTTCTCAAGGTGATATTTGTA 725
 Qy 850 AAAAAGAAATTTCTGTATCAGACCAAGATTTCTCAGACAACTTGTATTTACCTTTGGCC 909
 Db 726 AAAAAATGTTTCATGCCCACTACTTAGATCAATTTCTTGGTACGGAAGTGTGCTAC 785
 Qy 910 AGTGTATCTTGATCAGATTTGTAGTAATATATCATGTTACTTCTGGGTGGATTCACACCA 969
 Db 786 GGGAGCTGTAGATCTTCTCTGACGCAACGCTTTATTCATCTTCTGTGGATTTGACAAGA 845
 Qy 970 ACAATATGAACATGAGCCGAGCAAGTGTATATGCTGCCCACTCTTTGCTGGAACTCTG 1029
 Db 846 AAAAATTTAAATGTGAGTGGCTTTTGTATCTAGGGCATAATCCAGCAGGAACATCTA 905
 Qy 1030 TGCAAAATATTTCTACACTGGAGCCAGCAGTGAATTTCTGGTGAATCTCCGGGCATTTGACT 1089
 Db 906 CTCAGACCTTTTCCACTGGGCAAGCTTGTCTAAATCTGGAAAGCTTCAAGCCCTATACT 965
 Qy 1090 GGGGAGTGAACCAAAATCTGAAATAATGCAATCAGCCAACTCTCTGTAAAGTACAGAG 1149
 Db 966 GGGGAAGTCCATTACAGAAATGTTTACACTACATCAGAAAAAGCCCTCCCTACTATGATG 1025
 Qy 1150 TCAGAGATATCAGCGTCCCTACAGCAATGTGRCAGGAGTCAAGGCTGGCTTTCAAATC 1209
 Db 1026 TGTAGCCATGACCGTGGCAATTTGAGTGTGAAACGCTGGCCATGATCTCTGGCTGATC 1085
 Qy 1210 CAGAAGAGTGAATGCTCTCTGAGGTGACCAACCTCATCTACCATAGAAATTC 1269
 Db 1086 CCCAAGATGTGCAATGTGCTTCCCAAACTCCCAACCTTCTGTACCATAGGAGATTC 1145
 Qy 1270 CTGAATGGCTCAGTGGATTTTCATCTGGGTTTGGATGCTCTCCCTACCGTATGTACAATG 1329
 Db 1146 TTCCCTACAATCAGCTGGACTTTCATCTGGGCGATGGATGGCCCTCAGGAGTTTACATG 1205
 Qy 1330 AAATCATCCATCTGATG 1346
 Db 1206 AGATAGTTACCATGATG 1222

Search completed: February 19, 2004, 10:30:12
 Job time : 4478.52 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 04:49:02 ; Search time 126.478 Seconds
(without alignments)
7133.182 Million cell updates/sec

Title: US-10-042-431-45
Perfect score: 2044
Sequence: 1 gtgacccacgctccggg.....aaaaaaaaaaggcgccgc 2044

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgm2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgm2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgm2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgm2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgm2_6/ptodata/2/ina/PTCUS_COMB.seq.*
6: /cgm2_6/ptodata/2/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	470	23.0	1367	4	US-09-348-930A-8
2	456.8	22.3	1411	4	US-09-186-489-3
3	456.8	22.3	1411	4	US-10-043-665B-3
4	453.6	22.2	1134	4	US-09-186-489-4
5	453.6	22.2	1134	4	US-10-043-665B-4
6	450.4	22.0	1134	4	US-09-186-489-1
7	450.4	22.0	1134	4	US-10-043-665B-1
8	449.8	22.0	8324	4	US-09-186-489-5
9	449.8	22.0	8324	4	US-10-043-665B-5
10	444.2	21.7	1137	1	US-08-227-108-2
11	444.2	21.7	1137	2	US-09-073-674-2
12	444.2	21.7	1140	1	US-08-227-108-4
13	444.2	21.7	1140	2	US-09-073-674-4
14	444.2	21.7	1146	1	US-08-227-108-6
15	444.2	21.7	1146	2	US-09-073-674-6
16	444.2	21.7	1198	4	US-09-348-930A-5
17	444.2	21.7	1528	4	US-08-227-108-1
18	444.2	21.7	1528	2	US-09-073-674-1
19	444.2	21.7	1528	4	US-09-348-930A-1
20	444.2	21.7	1528	4	US-09-348-930A-7
21	414.2	20.3	1360	4	US-09-820-001-1
22	388.8	19.0	1048	4	US-09-348-930A-3
23	104.2	5.1	22067	4	US-09-820-001-3
24	48.2	2.4	178	1	US-08-751-782-2
25	48.2	2.4	178	2	US-08-751-782-1
26	46.4	2.3	43676	3	US-08-925-171-2
27	42	2.1	8041	1	US-08-765-081-1

C	28	42	2.1	8041	3	US-09-098-082-1	Sequence 1, Appli
C	29	42	2.1	8041	5	PCT-US95-06994-1	Sequence 1, Appli
	30	41.4	2.0	87563	4	US-09-453-702B-57	Sequence 57, Appli
	31	40.4	2.0	1140	2	US-08-698-805-5	Sequence 5, Appli
	32	26000	1.9	26000	4	US-09-843-376-10	Sequence 10, Appli
	33	39.8	1.9	168575	4	US-09-426-290-1	Sequence 1, Appli
	34	39.4	1.9	53332	4	US-09-801-861-3	Sequence 3, Appli
C	35	39	1.9	874	4	US-09-389-681-426	Sequence 426, App
C	36	39	1.9	874	4	US-09-620-405B-426	Sequence 426, App
C	37	39	1.9	874	4	US-09-433-826B-426	Sequence 426, App
C	38	39	1.9	874	4	US-09-604-287A-426	Sequence 426, App
	39	38.6	1.9	2815	1	US-08-230-491A-1	Sequence 1, Appli
	40	38.6	1.9	2815	1	US-08-619-280A-1	Sequence 1, Appli
	41	38.6	1.9	2815	2	US-08-940-391-1	Sequence 1, Appli
C	42	38.6	1.9	1664976	4	US-08-916-421B-1	Sequence 1, Appli
	43	38.4	1.9	2798	3	US-09-318-448-34	Sequence 34, App
	44	38	1.9	215	4	US-09-702-705-499	Sequence 499, App
	45	38	1.9	215	4	US-09-736-457-499	Sequence 499, App

ALIGNMENTS

RESULT 1

US-09-348-930A-8
Sequence 8, Application US/09348930A
Patent No. 6573431
GENERAL INFORMATION:
APPLICANT: Lence, P
APPLICANT: Gruber, V
APPLICANT: Baudino, S
APPLICANT: Merot, B
APPLICANT: Benicourt, C
APPLICANT: Cudrey, C
TITLE OF INVENTION: Recombinant Predutodenal Lipases and Polypeptides Derivatives Produced by Plants, Processes for Obtaining Them and Their Uses
FILE REFERENCE: 18433/2012
CURRENT APPLICATION NUMBER: US/09/348,930A
CURRENT FILING DATE: 1999-07-02
PRIOR APPLICATION NUMBER: 08/945,321
PRIOR FILING DATE: 1998-02-12
PRIOR APPLICATION NUMBER: PCT/FR96/00606
PRIOR FILING DATE: 1996-03-19
NUMBER OF SEQ ID NOS: 37
SOFTWARE: PatentIn version 3.0
SEQ ID NO 8
LENGTH: 1367
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (47)..(1240)
US-09-348-930A-8

Query Match	23.0%	Score 470;	DB 4;	Length 1367;
Best Local Similarity	64.2%	Pred. No. 1.4e-124;	Indels	0;
Matches	707;	Conservative	0;	Mismatches 395;
				Gaps 0;
Qy	244	CTGTGACCCAGAGCATTTCATGAATATTAGTGAATAATCAATCCAAATCAAGGCTATCCCT	303	
Db	123	CTGGAAGCCCTGAAGTGACTATGAACATTAGTCAGATGATTACTTATTGGGGATACCCAA	182	
Qy	304	GTGAGAAATATGAATGCGCACTGAGATGGGTATCTCTTCTTAAACAGATTCTTC	363	
Db	183	ATGAAGAAATATGAATGTTGATGAGATGGTTATTCTTGAAGTCAATGAATTCCTT	242	
Qy	364	GAGGCTTAGTGCACACCTAAAGAGACAGAGTTCAGGCTGTGGTGTGTACTTCGACGATGGCC	423	
Db	243	ATGGGAAGAAAAATTCAGGGAATACAGGCCAGAGACCTGTGTGTTTTTTCAGCATGGTT	302	
Qy	424	TAGTTCGAGGTCTAGCAACTGGATTTCACCTGCCCAACAAATAGCCTGGGCTTCATTC	483	
Db	303	TGCTTGCATCAGCCCAAACTGGATTTCACCTGCCCAACAACTTCGCTTCATTC	362	

484 TGGCAGATGCTGGTGTGATGCTGGATGGGAAACAGCAGGGGAAACGCTGCTCGAA 543
Db TGGCAGATGCTGGTGTGATGCTGGATGGGAAACAGCAGGGGAAACGCTGCTCGAA 422
QY 544 AACACAGACACTCTCCATAGACCAAGATGAGTCTGGGCTTTTCAGTTATGATGAGTGG 603
Db 423 GAAACTGTACTATTACACAGATTCAGTTGAATCTCGGCTTTTCAGTTATGATGAGTGG 482
QY 604 CTAGGTTGACCTTCTGAGTGAATTAATTTTTCAGAAACAGGCGGCGGCGGAAAGA 663
Db 483 CTAATATGACCTTCCAGCCCAATCACTTCAATGTAAGAAACAGGCGGCGGCGGCGGCGG 542
QY 664 TCTATTATGCTGGCTATTCACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 723
Db 543 TACACTATGTTGGCCATTCACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 602
QY 724 CAGAGCTGGCTCAGAAATCAAAATGATTTTTCAGTTCAGCAGCAGCAGCAGCAGCAGC 783
Db 603 CCAGCCTGGCTAAAGAAATCAAAACCTTCTATGCTCTAGCTCTGCTGCTGCTGCTGCT 662
QY 784 ATGCAAAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 843
Db 663 ATACAAAGGCTTATACAAACTTGAATTTCTTCTCTCAATCCCTCTTCAAGTTATAT 722
QY 844 TTGGCAAAAGGAAATTTCTGATTCAGACAGATTTCTCAGACAACTTTGTTATTTACCTTT 903
Db 723 TTGGTGACAAATATTTCTACCCACACAACTTCTTGTATCAATTTCTTGTCTACTGAAGT 782
QY 904 GTGGCCAGGCTGATTTCTGATCAGATTTGATGATGATGATGATGATGATGATGATGAT 963
Db 783 GCTCCGCTGAGTGTGATCTCTCTTGGCAGCAATGCTTATTAATTTGATTTGATTTG 842
QY 964 ACACCAACATATGAACATGAGCGGAGCAAGTGTATGCTGCGCCACACTCTTTGCTGGAA 1023
Db 843 ACAGTAAAGACTTTAAACAGGCTGCTTGGATGTGATCTATCACATAATCCAGCAGGAA 902
QY 1024 CATCTGTCAGAAATATTTCTACCTGAGCGAGCAGTGAATCTGCTGAGTCTGCTGAGT 1083
Db 903 CTTCCTGTTCAACATGCTTCCATGACCGGCTGTTAGTCTGGAATTTCCAGCTT 962
QY 1084 TTGACTGGGGAGTGAGACCAAAATCTGGAAATATGCAATCAGCCAACTCTCTGTAAGT 1143
Db 963 ATGACTGGGGAAGCCAGTTTCAAGATAGGATGCACTATGATCAGTCCCACTCTCTACT 1022
QY 1144 ACAGGTCAGATATGACGCTCCCTACCAATGTCAGCAGGCTCAGGCTGCTGCTTT 1203
Db 1023 ACAATGTGAGCCATGAATGTACCAATGTCAGTGTGGAACGGTGGCAAGGACCTGTTGG 1082
QY 1204 CAAATCCAGAGAGCTGAAAATGCTCTCTGAGGTGACCAACCTCATCTACCAATAGA 1263
Db 1083 CTGACCCCAAGATGTTGGCTTTTCTTCCAAACTCCCAATCTTATTTTACCAAGG 1142
QY 1264 ATATTCCTGATGGCTCAGTGAATTTCTCTGGGTTGATGCTCTCTCAGCGTATGT 1323
Db 1143 AGATTCTCTTTTACAATCAGTGAATTTCTCTGGGCAATGATGCTGCTGCTGCTGCT 1202
QY 1324 ACAATGAATCATCCATCTGAT 1345
Db 1203 ACAATGACATTTTCTATGAT 1224

RESULT 2

US-09-186-489-3
; Sequence 3, Application US/09186489
; Patent No. 6375947
; GENERAL INFORMATION:
; APPLICANT: Bolen, Paul L.
; APPLICANT: Cihak, Paul L.
; APPLICANT: Scharpf Jr., Lewis G.
; TITLE OF INVENTION: Purified Recombinant Kid Pregastric Esterase, and
; TITLE OF INVENTION: Processes for its Production and Use
; FILE REFERENCE: 5499/3

; CURRENT APPLICATION NUMBER: US/09/186,489

; CURRENT FILING DATE: 1998-11-05

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3

; LENGTH: 1411

; TYPE: DNA

; ORGANISM: Kid (Goat)

; US-09-186-489-3

Query Match 22.3%; Score 456.8; DB 4; Length 1411;

Best Local Similarity 62.0%; Pred. No. 8.4e-121;

Matches 741; Conservative 0; Mismatches 452; Indels 3; Gaps 1;

QY 167 AATGAAATCTGGCTTCTGATTTCTGGTGGGTATATGTTCCAGAGAAATGTAATTCAGT 226

Db 61 ATTGAAATCTGGTGGCTTCTGTAACGGTGTGTTTCATCCACATGCTGGAATGCATT 120

QY 227 ACATATGC---CAACTAAGCTGTGACCCAGAGCAATTCATGAATATTAGTGAATCAT 283

Db 121 TTGTTTCTTGGAAATTTGCTAAGAACCCCTGAAGCCAGTATGAATGTGAGTCAGATGAT 180

QY 284 CCAACATCAAGGCTATCCCTGTGAGGAATATGAAGTCGCAACTGAAGATGGGTATATCCT 343

Db 181 TTCTTCTGGGCTTACCAAGTGAATGATGATGATGATGATGATGATGATGATGATGAT 240

QY 344 TTCTGTAAACAGGATTCCTCGAGGCTGATGCAACCTAAGAACAGAGTTCAGAGGCTGT 403

Db 241 TCAGGTCTATCGGATTCCTCATGGAAGATGATGCTTAATCATTTAGGTGAGAGACCTGT 300

QY 404 GGTCTTATCGAGCATGGCTTAGTTGGAGTGTAGCAACTGATTTCCAACTGCCCAA 463

Db 301 TGTGTTTCTGAGCATGGCTCTTCTTGGCTCAGCTACAACTGGATTTCCAACTTCCCAA 360

QY 464 CAATAGCTGGCTTCTTCTGAGCATGCTGTTTGTGAGCTGTGAGTGGGAAACAGCAG 523

Db 361 CAACAGCTGGGCTTCTCTGAGCATGCTGTTTGTGAGCTGTGAGTGGGAAACAGCAG 420

QY 524 GGGAAACGCTGCTCGAAACACAGACACTCTCATAGACCAAGATGATGATGATGATGAT 583

Db 421 AGGAAACACTTGGGCTCGAGCAATTTATATCTTATCTTATCTTATCTTATCTTATCT 480

QY 584 TTTCAAGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 643

Db 481 TTTCAAGTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540

QY 644 GAAACGCGGCGGAGGAAAGATCTATTATGCTGCTTATGCTGCTTATGCTGCTTATGCT 703

Db 541 GAGAACAGGACAGAGAGGCTACACTATGTTGGCCATTTCCAGGCGGCGGCGGCGGCGG 600

QY 704 TATTGCAATTTTCCACCATGCGCAGAGCTGGCTCGAAATCAAAATGATTTTGTCTTATG 763

Db 601 TGTGCTTTTCTTACCAATCCCACTGGCTGAAATATGACCTTCCATCTACAACTGATTT 660

QY 764 ACCCATGACCTGTTAAGCATGCAAAAGCCCGGAGCAAAATTTTGTGCTGCGCAGA 823

Db 661 CCCAGTCCGACAGTGAAGCACACCCAGAGCTGTTTAAACAACTTGCATTTATTCCTCA 720

QY 824 TATGATGATCAAGGATTTGTTGGCAAAAGAAATTTCTGTATCAGACCAAGATTTCTCAG 883

Db 721 CTTCTCTTCTCAAGATTATTTTGTACAAATGTTCTTACCCACACAAATTTTGTGAACA 780

QY 884 ACACTTGTATTACCTTTTGTGCGCAGGTGATTTCTTGTATCAGATTTTGTATGATATCAT 943

Db 781 ATTTCTTGGTGTGAAAGTGTCTCTCGTGAGACACTGATGATGATGATGATGATGATGAT 840

QY 944 GTTACTTCTGGTGGATTCACACCAATATGAATGAGTGGCGGAGCAAGTGTATATGC 1003

Db 841 GTTGTCCATTTACTGGAGCTGACATTAATACTTCAACATGATGCTGTAGATGTATGT 900

QY 1004 TGCCCACTCTTGTGGAACATCTGTGCAAAATTTCTCACTGGAGCGGAGGAGTGA 1063

Db 901 AGCACATAATCCAGCAGGAGCTTCTGTTTCAAAACATCTCTCCACTGGAGAGAGGCTAT 960

QY 1064 TTCTGGTGAATCCGGGCAATTTGACTGGGAGTGAGACCAAAATCTGGAATAATGCAA 1123
 Db 961 GTCTGGGAAATTCGAAGCTTTTGACTGGGAGCCTCAGTTGAGAACCTTAATGCAATTATA 1020
 QY 1124 TCAGCCAACTCTGTGAAGGTACAGAGTCAGAGATATGACGGTCCCTACAGCAATGTGGAC 1183
 Db 1021 TCAGCCCAACACTCCCATCTCAATTTAAACAGCATGAATGTCCCAATTCAGATGAG 1080
 QY 1184 AGGAGTCAGACCTGCTTTCAATCCAGAGAGAGTGAAATGCTCTCTGAGGTGAC 1243
 Db 1081 TGCTGGCCCAAGACCTGTTGGGTGACCTCCAGATGTTGACCTTTTGTCTTCAAACTCTC 1140
 QY 1244 CAACCTCATCTACCAATGAATATTTCTGAAATGGGCTCACGTGGATTTTCACTGGGGTTT 1303
 Db 1141 TAATCTCATTCACCAAGGAATTTCCAAATTAACATCATCTGCACTTTATCTGGGCAAT 1200
 QY 1304 GGATGCTCTCACCGTATGTAATGAATCATCCATCTGATGAGCAGAGGAGA 1359
 Db 1201 GGATGCACTCAAGAGTTTACAATGAATTTATTTCTTGATGCAAAAGACAAA 1256

RESULT 3

US-10-043-665B-3
 ; Sequence 3, Application US/10043665B
 ; Patent No. 6582948
 ; GENERAL INFORMATION:
 ; APPLICANT: Bolen, Paul L.
 ; APPLICANT: Cihak, Paul L.
 ; APPLICANT: Scharpf Jr., Lewis G.
 ; TITLE OF INVENTION: Recombinant Kid Pregastric Esterase and Methods for Its
 ; TITLE OF INVENTION: Production and Use
 ; FILE REFERENCE: IFF-0009
 ; CURRENT APPLICATION NUMBER: US/10/043.665B
 ; CURRENT FILING DATE: 2002-09-24
 ; PRIOR APPLICATION NUMBER: US 09/186,489
 ; PRIOR FILING DATE: 1998-11-05
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 1411
 ; TYPE: DNA
 ; ORGANISM: Kid (Goat)
 US-10-043-665B-3

Query Match 22.3%; Score 456.8; DB 4; Length 1411;
 Best Local Similarity 62.0%; Pred. No. 8.4e-121;
 Matches 741; Conservative 0; Mismatches 452; Indels 3; Gaps 1;
 QY 167 AATGAAATGTGGCTTCTGATTTCTGGTGGCGTATATGTTCCAGAGAAATGTGAATTCAGT 226
 Db 61 ATTGAATGTGGTGGCTTCTGATTTCTGGTGGCGTATATGTTCCAGAGAAATGTGAATTCAGT 120
 QY 227 ACATATGC---CACTAAAGCTGTGGACCCAGAGAGCATTCATGAATATTAGTGAATCAT 283
 Db 121 TTGTTTCTTGGAAAAATTTGTAAGAACCTTGAAGCCAGCATGTAATGTGAGTCAGATGAT 180
 QY 284 CCAACATCAAGCTATCCCTGTGAGGAATATGAAGTGGCACTCAAGATGGGTATATCCT 343
 Db 181 TTCTTCTGGGGCTACCCAGTGAATGATTAAGTGTATTAATGCAATGCTGATATATCCT 240
 QY 344 TTCTGTTAAAGGATTCCTCGAGGCTTAGTGCAACCTTAAGAGACAGGTTCCAGGCTGT 403
 Db 241 TCAGGTCTATCGGATTCCTCATGGAAGAAATGATGCTTAATCATTTAGGTGAGAGACCTGT 300
 QY 404 GGTGTTACTGAGCATGCGCTAGTTGGAGGTGCTAGCACTGGATTTCCAACTGGCCCA 463
 Db 301 TGTGTTCTGAGCAATGCTTCTTGCTGAGTACAACTGGAATTTCCAACTTCCCAA 360
 QY 464 CAATAGCTGGCTTCAATCTTGGCAGATGCTGTTTTCAGTGTGAGTGGGGAACAGCAG 523
 Db 361 CAACAGCTGGCTTCTCTCTGGCAGATGCTGTTTATGATGCTGTTGGTGGGAACAGCAG 420

QY 524 GGGAAACGCTCTGCTCGAATAACACAGACACTCTCCATAGACCAGATGAGTTTCTGGC 583
 Db 421 AGGAAACACTTTGGGCCCCAGGAAATTTATATTAATTCACAGACTCCCTGAAATTTGGC 480
 QY 584 TTTTCAGTTATGATGAGATGGCTTAGCTTTGACCTTCTCTGAGTGATAAACTTTATTTTGA 643
 Db 481 TTTTCAGTTTATGATGAATGGCTGAATATGACCTTCCATCTACAAATGATTTTCTTAA 540
 QY 644 GAAACGGCCACAGGAAGATCTATATGTGGCTATTTCAGGGCAGCACCATGGGTT 703
 Db 541 GAAACAGACAGAGAGAGCTACACTATGTTGGCCATTTCCCAAGGCAACCAATTTGGTT 600
 QY 704 TATTGCAATTTCCACATGCCAGAGCTGGCTCAGAAAAATCAAAATGATTTTGTCTTTAGC 763
 Db 601 TGTGCGCTTTTCTACCAATCCACACTGGCTGGTGAATAATCGAGTCTTCCATGCAATGAG 660
 QY 764 ACCATAGCCACTGTTTAGCATGCAAAAGCCCCGGGACCAAAATTTTGTGTGCCAGA 823
 Db 661 CCCAGTGGCCACAGTGAAGCACACCCAGAGCCTGTTTAAACAAACTTGCATTTATCTCA 720
 QY 824 TATGATGATCAAGGCAATTTGTTGGCAAAAGAAATTTCTGTATCAGACAGATTTCTCAG 883
 Db 721 CTTCCTCTTCAGATATATTTGGTAAACAAATGTTCTACCCACAAATTTTGTGAACA 780
 QY 884 ACAACTTGTATTATACCTTTTGGCCAGGTGATTTCTGATCAGATTTGTAGTAATCAT 943
 Db 781 ATTTCTTGGTGTGAAGTGTCTCTCTGAGACACTGGATGTCCTTTTAAAGATGCTT 840
 QY 944 GTTACTTCTGGGTGATTTCAACACCAACAATATGAACATGAGCCGAGCAGATGATATGC 1003
 Db 841 GTTGGCAATTAATGAGCTGCAATPAAAACTTCAACATGAGTCTGCTTGAATGTATGT 900
 QY 1004 TGCCCACTCTTGTGGAACATCTGTGCAAAATATTCTACCTGAGCCAGGCAAGTCAA 1063
 Db 901 AGCATAATATCCAGCAGGAGCTTCTGTTCAAAACATCTCTCAGTGAAGAGCTATTAA 960
 QY 1064 TTCTGTGAACTCCGGGCAATTTGACTGGGGAGTGAGACCAAAATCTGGAATAATGCAA 1123
 Db 961 GTCTGGAAATTTCCAGCTTTTGTGCTGGGAGCTCAGTGAACCTTAATGCAATATA 1020
 QY 1124 TCAGCCAACTCTGTAAAGTACAGAGTCAGAGATATGACGGTCCCTACAGCAATGTGAC 1183
 Db 1021 TCAGCCCAACACTCCATCTACAAATTTAAACAGCCATGAATGTCCCAATTTGCAATGAG 1080
 QY 1184 AGAGGTGAGACTGGCTTTCAATCCAGAGAGCTGAAATGCTGCTCTCTGAGGTGAC 1243
 Db 1081 TGCTGGCCCAAGACCTGTTGGGTGACCTCAGAGTGTGACCTTTGCTTTCAAACTCTC 1140
 QY 1244 CAACCTCATCTACCAATAAGAAATTTCTGAAATGGGCTCACGTGGATTTTCACTTGGGTT 1303
 Db 1141 TAACTCTCATTCACCAAGGAATTTCCAAATTAACATCATCTGAGCTTTATCTGGCAAT 1200
 QY 1304 GGATGCTCTCACCTATGTACATGAATCAATCATCTGATGAGCAGGAGGAGA 1359
 Db 1201 GGATGCACTCAAGAGTTTACAATGAATTTATTTCTTGATGCAAAAGACAAA 1256

RESULT 4

US-09-186-489-4
 ; Sequence 4, Application US/09186489
 ; Patent No. 6375947
 ; GENERAL INFORMATION:
 ; APPLICANT: Bolen, Paul L.
 ; APPLICANT: Cihak, Paul L.
 ; APPLICANT: Scharpf Jr., Lewis G.
 ; TITLE OF INVENTION: Purified Recombinant Kid Pregastric Esterase, and
 ; TITLE OF INVENTION: Processes for Its Production and Use
 ; FILE REFERENCE: 5499/3
 ; CURRENT APPLICATION NUMBER: US/09/186,489
 ; CURRENT FILING DATE: 1998-11-05
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 4

```

; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Bovine
US-09-186-489-4

Query Match      22.2%; Score 453.6; DB 4; Length 1134;
Best Local Similarity 62.7%; Pred. No. 6.1e-120;
Matches 705; Conservative 0; Mismatches 419; Indels 0; Gaps 0;

QY 236 AACTAAGCTGTGGAGCCAGAGCAATCATGAATATTAGTGAATCATCAATCAATCAAGG 295
Db 9 AAAAAATGTGAAGACCCCTGAAGCCAGTATGAATGTAGTCAGATGATTTCTACTGGGG 68

QY 296 CTATCCCTGTGAGGAATATGAAGTCCCACTGAAGATGGTATATCTTTCTGTTAAACAG 355
Db 69 CTACCAAGTGAGATGCATTAAGTTATATCTGGGATGGTATATCTTCAAGTCTATCG 128

QY 356 GATTCTCGAGGCTGTAGTGAACCTAAGAGACAGGTTCCAGGCTGTGGTGTACTGCA 415
Db 129 GATTCTCATGGAAGAATAATCTAATCATTTAGGTCTAGACACCTGTGTTTCTGCA 188

QY 416 GCATGGCTAGTTGGAGGCTGTAGCAACTGAGTATTCACCTGCCCAACATAGCTGGG 475
Db 189 GCATGGCTCTTCTTGGATCAGCCCAAACTGGATTTCCAACTGCCCAAGAACGCTGGG 248

QY 476 CTTCATTCTGCGAGATCTGTTTTCAGCTGTGGTGGGAAACAGCAGGGGAAACGCTG 535
Db 249 CTTCCTCTGGCAGATCTGTTTATGACGTGTGGTGGGAAACAGCAGAGGAACACCTG 308

QY 536 GTCTCGAAAAACACAAGACACTCTCCATAGAACCAAGATGATTTCTGGGCTTTTCAATTGA 595
Db 309 GGCCCGAGGAATTTATCTATTATTCACCAAGCTCCCGGAATCTGGGCTTTTCAAGTCTTGA 368

QY 596 TGAGATGGCTAGTTGACCTTCTGAGTGAATTAATTTTTCAGAAAAACGGGCA 655
Db 369 TGAATGGCGGAATATGACCTTCACTCAATGATTTTCACTTAAGGAGAACAGGACA 428

QY 656 GGAAGAGATCTATTATGTCGGCTATTTCAGAGGACCAACCATGGCTTTTATTCATTTTC 715
Db 429 GAAGAAGCTACATATGTTGGCCATTCCCAAGGACCAACCATGGCTTTTATTCGCTTTTC 488

QY 716 CACCATGCCAGAGCTGCTCAGAAAAATCAAAATGATTTTGTCTTTAGCACCACATGCCAC 775
Db 489 TACCAGTCCACATGTGCTGAAATAATCAAGTCTTCTATGCAATAGCCCGAGTTGCCAC 548

QY 776 TGTTAGCATGCAAAAGCCCGGAGCAAAATTTTGTGTCGCCAGATATGATGATCAA 835
Db 549 AGTGAAGTACCAAGAGGCTGTTTAAACAACTTGCACTTATTCCTCACTTCTCTCTCAA 608

QY 836 GGGATTTGTTGGCAAAAGAAATTTCTGTATCAGACAGATTTCTCAGACAACTTGTAT 895
Db 609 GATTATATTTGGTGAACAATGTTCTACCCACACTTTTGGAAATTTCTTGGTGT 668

QY 896 TTACCTTTGGCCAGGTGATTTTGAATCAGATTTGTATGTAATATCATGTTACTTCTGGG 955
Db 669 TGAATGTGTCCTGTCAGACACTGGATGTCCTTTTGAAGAATGCTTGTGGCCATTAC 728

QY 956 TGATTTCAACACCAATATGAATGACGCGAGCAAGTATATGCTGCCCACTCT 1015
Db 729 TGGATTTGACATATAAACTTCAACATGATGCTCTTGAATGATATATGACATATATCC 788

QY 1016 TGCTGGAACATCTGTGCAAAATATTCTACACTGGAGCCAGGAGTGAATTTCTGGTGAAT 1075
Db 789 AGCAGGAATCTCTGTTCAAAACACCTCCACTGGAGACAGGCTGTTTAACTCTGGAAAT 848

QY 1076 CCGGGCATTTGACTGGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCCACTCC 1135
Db 849 CCAAGCTTTTGACTGGGGAGCCCATATCAGAAACCTTAATGCAATATCATCAGCCACACC 908

QY 1136 TGTAAGGTACAGAGTACAGATATGACGGTCCCTCAGCAATATGTCAGAGGAGGTGAGGA 1195
Db 909 TCCCATCTACAAATTTAAACAGCCATGAATGTCCCAATTTGCAAGTATGGAGTGTGACATGA 968

```

```

RESULT 5
US-10-043-665B-4
; Sequence 4, Application US/10043665B
; Patent No. 6582948
; GENERAL INFORMATION:
; APPLICANT: Bolen, Paul L.
; APPLICANT: Cihak, Paul, L.
; APPLICANT: Scharf Jr., Lewis G.
; TITLE OF INVENTION: Recombinant Kid Pregastric Esterase and Methods for Its
; TITLE OF INVENTION: Production and Use
; FILE REFERENCE: IFF-0009
; CURRENT APPLICATION NUMBER: US/10/043.665B
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 09/186,489
; PRIOR FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: bovine
US-10-043-665B-4

```

```

Query Match      22.2%; Score 453.6; DB 4; Length 1134;
Best Local Similarity 62.7%; Pred. No. 6.1e-120;
Matches 705; Conservative 0; Mismatches 419; Indels 0; Gaps 0;

QY 236 AACTAAGCTGTGGAGCCAGAGCAATCATGAATATTAGTGAATCATCAATCAATCAAGG 295
Db 9 AAAAAATGTGAAGACCCCTGAAGCCAGTATGAATGTAGTCAGATGATTTCTACTGGGG 68

QY 296 CTATCCCTGTGAGGAATATGAAGTCCCACTGAAGATGGTATATCTTTCTGTTAAACAG 355
Db 69 CTACCAAGTGAGATGCATTAAGTTATATCTGGGATGGTATATCTTCAAGTCTATCG 128

QY 356 GATTCTCTGAGGCTGTAGTGAACCTAAGAGACAGGTTCCAGGCTGTGGTGTACTGCA 415
Db 129 GATTCTCATGGAAGAATAATGCTAATCATTTAGGTCTAGACACCTGTGTTTCTGCA 188

QY 416 GCATGGCTAGTTGGAGGCTGTAGCAACTGAGTATTCACCTGCCCAACATAGCTGGG 475
Db 189 GCATGGCTCTTCTTGGATCAGCCCAAACTGGATTTCCAACTGCCCAAGAACGCTGGG 248

QY 476 CTTCATTCTGCGAGATCTGTTTTCAGCTGTGGTGGGAAACAGCAGGGGAAACGCTG 535
Db 249 CTTCCTCTGGCAGATCTGTTTATGACGTGTGGTGGGAAACAGCAGAGGAACACCTG 308

QY 536 GTCTCGAAAAACACAAGACACTCTCCATAGAACCAAGATGATTTCTGGGCTTTTCAATTGA 595
Db 309 GGCCCGAGGAATTTATCTATTATTCACCAAGCTCCCGGAATCTGGGCTTTTCAAGTCTTGA 368

QY 596 TGAGATGGCTAGTTGACCTTCTGAGTGAATTAATTTTTCAGAAAAACGGGCA 655
Db 369 TGAATGGCGGAATATGACCTTCACTCAATGATTTTCACTTAAGGAGAACAGGACA 428

QY 656 GGAAGAGATCTATTATGTCGGCTATTTCAGAGGACCAACCATGGCTTTTATTCATTTTC 715
Db 429 GAAGAAGCTACATATGTTGGCCATTCCCAAGGACCAACCATGGCTTTTATTCGCTTTTC 488

QY 716 CACCATGCCAGAGCTGCTCAGAAAAATCAAAATGATTTTGTCTTTAGCACCACATGCCAC 775

```

```

Db      489  TACCAGTCCACATTGGCTGAAAAAATCAAGTCTTCTATGATAGCCCGAGTGGCCAC 548
QY      776  TGTAAAGATGAAAAAGCCCGGACCAAAATTTTGTGTCGCGAGATATGATGATCAAA 835
Db      549  AGTAAAGTACACAGAGGCGCTGTTAAACAACATGACATTAATCTCTCACTTCTCTTCAA 608
QY      836  GGGATTGTTGGCAAAAAAGAAATTTCTGTATCAGACGAGATTTCTCAGACAACATTTGTTAT 895
Db      609  GATTATATTGGTGACAAATGTTCTACCCACACATTTTGGAAACATTTCTTGGTGT 568
QY      896  TTACCTTTGGCGCAGGTGATTTCTGATCAGATTTCTAGTAATATCATCTTACTTTCTGGG 955
Db      669  TGAATATGTCGCGTGAGACATGCGATGTCCTTTGTAAGAATGCCCTTTTGGCCATTAC 728
QY      956  TGGATTCAACACCAACAAATATGAATGAGCGGAGCAAGTATATGCTGCGCCACACACT 1015
Db      729  TGGAGTTGACATTAATAAACTTCAACATGAGTCTGATGATGTTATATAGCAATAATCC 788
QY      1016  TGTGGAACATCTGTGCAAAATATTCTACACTGGAGCCAGGAGTGAATTTCTGGTGAAT 1075
Db      789  AGCAGGAATCTTGTGTTCAAAACACCCCTCCACTGGAGACAGGCTGTTAAGTCTGGAAAT 848
QY      1076  CCGGCATTTGACTGGGGAGTGCAGACCAAAATCTGGAATAATGCAATCAGCCCACTCC 1135
Db      849  CCAAGCTTTTGAATGGGGAGCCCTCAGTTGAGAACCTTAATGCATTAATCAGCCCACTCC 908
QY      1136  TGTAAAGTACAGAGTACAGATATAGCGTCCCTACAGCAATGTGAGAGGAGGTGAGGA 1195
Db      909  TCCCATCTCAATTTAAACAGCCATGAATGTCCCAATTCAGATGAGTGGAGTGCAGCAATGA 968
QY      1196  CTGGCTTTCAATCCAGAGACGTGAAATGCTCTCTGAGGTGACCACTCATCTA 1255
Db      969  CTGTTGGCTGACCCCTCAGGATGTTGACCTTTTGTCTTTCAAAACTCTCTAATCTCATTTCA 1028
QY      1256  CCATPAAGAAATTTCTGGAATGGGCTCAGCTGGATTTTCACTCTGGGGTTTGGATCTCTCA 1315
Db      1029  CCACAAGAAATTTCAAATCATCTGGACTTTATCTGGGCAATGATGATGACCTCA 1088
QY      1316  CCGTATGTACATGAATCATCTCATCTGATGAGCAGGAGGAGA 1359

```

RESULT 6

```

US-09-186-489-1
; Sequence 1, Application US/09186489
; Patent No. 6375947
; GENERAL INFORMATION:
; APPLICANT: Bollen, Paul L
; APPLICANT: Cihak, Paul L
; APPLICANT: Scharpf Jr., Lewis G.
; TITLE OF INVENTION: Purified Recombinant Kid Pregastric Esterase, and
; TITLE OF INVENTION: Processes for its Production and Use
; FILE REFERENCE: 5499/3
; CURRENT APPLICATION NUMBER: US/09/186,489
; CURRENT FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Kid (Goat)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) . (1134)
US-09-186-489-1

```

```

Query Match      22.0%; Score 450.4; DB 4; Length 1134;
Best Local Similarity 62.5%; Pred. No. 5.1e-119;
Matches 703; Conservative 0; Mismatches 421; Indels 0; Gaps 0;
QY      236  AACTAAGCTGTGGACCCAGACCATTCATGATATATTAGTAATCATCCAAATCAAGG 295

```

```

Db      9    AAAAATTGCTAAGAACCTCGAGCCAGTATGAATGTGAGTCAGATGATTTCTTCTGGGG 68
QY      296  CTATCCCTGTGAGGAATATGAAGTCGCAACTGAAGATGGGTATATCTTTCTGTAAACAG 355
Db      69  CTACCCAAAGTGAGATGATAAAGTTATAAATGAGATGGCTATATCTTCAAGGTCTATCG 128
QY      356  GATTCTCTCGAGGCTTAGTGCAACCTAAGAAGACAGGTTCCAGGCTGTGGTGTACTGCA 415
Db      129  GATTCTCTCATGGAAGAATGATGCTAATCTTATAGTGCAGAGACCTGTTGTGTTCTGCA 188
QY      416  GCATGGCTAGTTGGAGGTGCTAGCAACTGGATTTCCAACTGCCCCCAACAATAGCCCTGG 475
Db      189  GCATGGCTCTTCTGCTCAGCTACAACTGGAATTTCCAACTTCCCAACAACAGCCCTGG 248
QY      476  CTTCAATTTCTGGCAGATCTGGTTTGGCTGTGGATGGGAAACAGCGGGAACCGCTG 535
Db      249  CTTCTCTCTGGCAGATCTGGTTTATGACGTGTGGGAAACAGCAGAGAGAAACACTTG 308
QY      536  GTCTCGAAAACACAAGACACTCTCCATAGACCAAGATGAGTCTGGGCTTTTCAGTTATGA 595
Db      309  GGGCCAGGAACATTTACTATTATCCAGACTCCCTGGAATTTCTGGCTTTTCAGCTTTGA 368
QY      596  TGAGATGGCTAGTTGACCTCTCTGCGAGTGAATAAATTTATTTTGCAGAAAAACGGGCCA 655
Db      369  TGAATGGCTGAATATGACCTTCCATCTACAATTTGATTTTCACTTAAAGAGAAACAGGACA 428
QY      656  GGAAGAAGATCTATTATCTCGCTATTTCACAGGCGACCACTGCTTTATTTTCATTTTC 715
Db      429  GAAGAAGCTACATATGTTGGCCATTTCCAAAGGACCAACCATTTGTTTGTGCTCTTTC 488
QY      716  CACCATCCAGAGCTGCTCAGAAAATCAAAATGTATTTTGTCTTTAGACCCCAATAGCCAC 775
Db      489  TACCAATCCACACATGCTGCTGAAAATTCGAAGTCTTCCATGATGATAGCCCGAGTCCAC 548
QY      776  TGTAAAGCATGCAAAAGACCCCGGACCAAAATTTTGTCTGCTGCGAGATATGATGATCAA 835
Db      549  AGTGAAGACACCCAGAGCTGTTTAAACAATTTGCACTTATTTCTCTCATCTTCTTCAA 608
QY      836  GGGATTTGTTGGCAAAAAAGAAATTTCTGTATCAGACCAAGATTTCTCAGACAACCTGTTAT 895
Db      609  GATTATATTTGGTAAACAAATGTTTACCCACACAAATTTTGTGAACAAATTTCTTGTGTGT 668
QY      896  TTACCTTTTGGCCAGGTGATTTCTGATCAGATTTGTAGTAATATCATGTTACTTCTGGG 955
Db      669  TGAAGTGTGCTCTGTCGAGACACTGGATGTCCTTTGTAAAGATGCTTGTGTCCTATAC 728
QY      956  TGGATTCAACACCAACAAATATGAACATGAGCCGAGCAAGTATATGCTGCCCCACACTCT 1015
Db      729  TGGAGCTGACAAATAAAGAACTTCAACATGAGTCTGTTAGTGTATGTAGCACATATCC 788
QY      1016  TGTGGAACATCTGTGCAAAATATTTCTAAGTGGAGCCAGGAGTGAATTTCTGGTGAAT 1075
Db      789  AGCAGGAGCTTCTGTTTCAAAACATCTCTCCACTGAGACAGGCTTATTAAGTCTGGGAAAT 848
QY      1076  CCGGGCATTTGACTGGGGAGTGCAGACCAAAATCTGGAATAATGCAATCAGCCCACTCC 1135
Db      849  CCAAGCTTTTGAATGGGGAGCCCTCAGTTGAGAACCTTAATGCATTAATCAGCCCACTCC 908
QY      1136  TGTAAAGTACAGAGTACAGATATAGCGTCCCTACAGCAATGTGAGAGGAGGTGAGGA 1195
Db      909  TCCCATCTCAAAATTTAAACAGCCATGAATGTCCCAATTCAGATGAGTGTGAGTGTGCGCAAGA 968
QY      1196  CTGGCTTTCAAAATCCAGAGACGTGAAATGCTCTCTGAGGTGACCACTCATCTA 1255
Db      969  CTGTTGGCTGACCCCTCAGGATGTTGACCTTTTGTCTTTCAAAACTCTCTAATCTCATTTCA 1028
QY      1256  CCATPAAGAAATTTCTGGAATGGGCTCAGCTGGATTTTCACTCTGGGGTTTGGATCTCTCA 1315
Db      1029  CCACAAGAAATTTCAAATCATCTGGACTTTATCTGGGCAATGATGATGACCTCA 1088
QY      1316  CCGTATGTACATGAATCATCTCATCTGATGAGCAGGAGGAGA 1359

```



```

Db 1089 AGAAGTTTACATGAATTAATTTCTTGTATGCGCAAGACAA 1132
US-10-043-665B-1
RESULT 7
; Sequence 1, Application US/10043665B
; Patent No. 6582948
; GENERAL INFORMATION:
; APPLICANT: Bolen, Paul L.
; APPLICANT: Cihak, Paul L.
; APPLICANT: Scharpf Jr., Lewis G.
; TITLE OF INVENTION: Recombinant Kid Pregastric Esterase and Methods for its
; TITLE OF INVENTION: Production and Use
; FILE REFERENCE: IFF-0009
; CURRENT APPLICATION NUMBER: US/10/043.665B
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 09/186,489
; PRIOR FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Kid (Goat)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1134)
US-10-043-665B-1
Query Match 22.0%; Score 450.4; DB 4; Length 1134;
Best Local Similarity 62.5%; Pred. No. 5.1e-119;
Matches 703; Conservative 0; Mismatches 421; Indels 0; Gaps 0;

Qy 236 AACTAAAGCTGTGGACCCAGCAAGCATTCATGAATATTAGTGAATCATCAACATCAAG 295
Db 9 AAAAAATTCCTAAGAACCTCGAAGCCAGTATGAATGTGAGTCAGATTTCTTCCTGGGG 68

Qy 296 CTATCCCTGTGAGGAATATGAAGTCGCAAGTGAAGTGGTATATCTTCTGTTAAACAG 355
Db 69 CTACCCAGTGAAGTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 128

Qy 356 GATTCTCTGAGGCTAGTGCACCTAAGAACAGAGGTTCCAGGCTGTGTTGTTACTGCA 415
Db 129 GATTCTCTGAGGCTAGTGCACCTAAGAACAGAGGTTCCAGGCTGTGTTGTTACTGCA 188

Qy 416 GCATGGCTAGTGGAGGTGTGCAACTGATTTCCAACTGCGCCCAACAAATAGCTGGG 475
Db 189 GCATGGCTAGTGGAGGTGTGCAACTGATTTCCAACTGCGCCCAACAAATAGCTGGG 248

Qy 476 CTTCATTTCTGCGAGATGCTGTTTGTGAGTGTGAGTGGGAAACAGAGGGAACGCTG 535
Db 249 CTTCCTCTGCGAGATGCTGTTTGTGAGTGTGAGTGGGAAACAGAGGGAACGCTG 308

Qy 536 GTCTCGAAACACAAAGACATCTCTCCATAGAACAGATGATTCGCGCTTTCAGTTATGA 595
Db 309 GCGCCGAGAACATTTATATCTATTCACCAAGTCTCCCTGAAATTCGCGCTTTCAGTTATGA 368

Qy 596 TGAGATGCTAGGTTGACCTTCTGAGTGATAAACTTTATTTGCGAAGAACGGGCCA 655
Db 369 TGAATGCTGAATGATGACCTTCTCATACATTTGATTTCTATTTAAGAGAACAGACA 428

Qy 656 GGAAGAAGCTATTATGTCGCTATTACAGGGCACCACCATGGGCTTTATTCATTTTC 715
Db 429 GAAGAAGCTACTATGTCGCTATTACAGGGCACCACCATGGGCTTTATTCATTTTC 488

Qy 716 CACCATGCGAGGTGGCTGAGAAATCAAAATGATTTTGTGCTTACGACCCATAGCCAC 775
Db 489 TACCATCCACACTGGCTGAGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 548

Qy 776 TGTAAAGCATCAAAAGCCCGGGACCAAAATTTTGTGCTGCGAGATGATGATCAA 835
Db 549 AGTGAAGCACACCCAGAGCCCTGTTTAAACAACTTGCACTTATCTCTCACTTCTCTTCAA 608

Query Match 22.0%; Score 449.8; DB 4; Length 8324;
Best Local Similarity 62.7%; Pred. No. 2.2e-118;
Matches 700; Conservative 0; Mismatches 417; Indels 0; Gaps 0;

Qy 243 GCTGTGGACCCAGAGCAATTCATGAATATTAGTGAATATCAACATCAAGGCTATCCC 302
Db 1547 GCTAAGAACCCCTGAAGCCAGTATGAATGTGAGTCAGATGATTTCTTCTGGGCTACCCA 1606

Qy 303 TGTGAGGAATATGAAGTCGCAACTGAGATGGGTATATCTTCTGTTAAAGATTTCT 362
Db 1607 AGTGAGTGCATAAAGTTATTAACATGCGATGGCTATATCTTCTGAGGTCTATCGGATTCCT 1666

Qy 363 CGAGGCTAGTGCACCTCAACACAGAGGTTCCAGGCTGTGTTGTTTACTGACAGATGGC 422
Db 1667 CATGGAAGATGATGCTAATCAATTTAGTCAAGACCTGTTGTTTCTGACAGATGGT 1726

```

```

RESULT 8
US-09-186-489-5
; Sequence 5, Application US/09186489
; Patent No. 6375947
; GENERAL INFORMATION:
; APPLICANT: Bolen, Paul L.
; APPLICANT: Cihak, Paul L.
; APPLICANT: Scharpf Jr., Lewis G.
; TITLE OF INVENTION: Purified Recombinant Kid Pregastric Esterase, and
; FILE REFERENCE: 5499/3
; CURRENT APPLICATION NUMBER: US/09/186.489
; CURRENT FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 8324
; TYPE: DNA
; ORGANISM: Yeast YE-1 expression vector
US-09-186-489-5

```

423 CTAGTTGAGTGTAGCACTGATTTCCAACTGCTCCCAACATAGCTGGCTTCATT 482
 1727 CTTCTTGCTCAGTACAACTGATTTCCAACTGCTCCCAACATAGCTGGCTTCCTC 1786
 483 CTGCGAGATGCTGTTTGGTGTGATGGGGAACAGCAGGGGAAAGCGCTGCTCGA 542
 1787 CTGCGAGATGCTGTTTGGTGTGATGGGGAACAGCAGGGGAAAGCGCTGCTCGA 1846
 543 AAACACAGACACTCTCCATAGACAAGATGATGCTGGGCTTCAGTTATGATGATG 602
 1847 GAACATTTATCTATTTACAGATCTCCCTGAAATCTGGGCTTCAGCTTTGATGAATG 1906
 603 GCTAGTTGACCTTCTGCTGAGTAACTTTATTTTGCAGAAACGGGCGGAGAAAG 662
 1907 GCTGAATATGACCTTCCATCTACAAATGATTTATCTTAAAGAGAACAGGACAGAG 1966
 663 ATCTATTATGCTGCTATTCAGGCGACCACTGCTGCTTTATGCTATTTCCACCATG 722
 1967 CTACACTATGCTGCTATTCAGGCGACCACTGCTGCTTTATGCTATTTCCACCAT 2026
 723 CCAGAGCTGCTCAGAAATCAAAATGATTTTCTTTAGCACCACATAGCACTGTTAAG 782
 2027 CCCACATGCTGAAATATGAAATGCTTCCATGATTTAGCCCGAGTCGCAAGTGA 2086
 783 CATGCAAAAGCCCGGAGCAAAATTTTGTGCTGCGAGATGATGATGATGATGATG 842
 2087 CACACCCAGAGCTGTTTAAACAACTTGCACTTATCTCTCACTTCTCTTCAAGATTAT 2146
 843 TTTGGCAAAAGAAATTTCTGATCAGACCAAGATTTCTCAGACAACTTGTATTACCTT 902
 2147 TTTGGTAAACAAATTTCTACCCACAAATTTTGTGAACTTCTTGTGTTGATG 2206
 903 TGTGCGAGGTGATTTGATGATGATTTGATGATTTGATGATTTGATGATTTGATG 962
 2207 TGCTCTGCTGAGACACTGATGCTTTTGTGAAGATGCTTTGCTGCTTCTTCAAGCT 2266
 963 AACACCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1022
 2267 GACATATAAACTTCACATGATGCTGCTGATGATGATGATGATGATGATGATGATG 2326
 1023 ACATCTGTGCAAAATATTTCTACATGAGCGAGCGAGTGAATTTCTGCTGAACTCCGGCA 1082
 2327 GCTTCTGTGTTCAAAACATCTCCATGAGGAGCAGGCTATTAAGTCTGGGAAATCCAAGCT 2386
 1083 TTTGACTGGGGAGTGGAGACCAAAATCTGGAATATGCAATGCAATGCAATGCAATG 1142
 2387 TTTGACTGGGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2446
 1143 TACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1202
 2447 TACAATTTACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2506
 1203 TCAATCCAGAGAGTGAATGCTGCTGCTGAGTGAACCTGATGATGATGATGATGATG 1262
 2507 GCTGACCTTCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2566
 1263 AATATTCCTGATGGGCTCAGTGGATTTCACTGCGGCTTTGATGATGATGATGATGATG 1322
 2567 GAAATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2626
 1323 TACATGAAATCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1386
 2627 TACAATGAAATTTATTTCTTGTGATGATGATGATGATGATGATGATGATGATGATG 2686

RESULT 9

US-10-043-665B-5
 ; Sequence 5, Application US/10043665B
 ; Patent No. 6582948
 ; GENERAL INFORMATION:
 ; APPLICANT: Bolen, Paul L.
 ; APPLICANT: Cihak, Paul, L.

; APPLICANT: Scharpf Jr., Lewis G.
 ; TITLE OF INVENTION: Recombinant Kid Pregastric Esterase and Methods for Its
 ; FILE OF INVENTION: Production and Use
 ; FILE REFERENCE: IFF-0009
 ; CURRENT APPLICATION NUMBER: US/10/043,665B
 ; PRIOR FILING DATE: 2002-09-24
 ; PRIOR APPLICATION NUMBER: US 09/186,489
 ; PRIOR FILING DATE: 1998-11-05
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 5
 ; LENGTH: 8324
 ; TYPE: DNA
 ; ORGANISM: Yeast YE-1 expression vector
 ; US-10-043-665B-5

Query Match 22.0%; Score 449.8; DB 4; Length 8324;
 Best Local Similarity 62.7%; Pred. No. 2.2e-118; Indels 0; Gaps 0;
 Matches 700; Conservative 0; Mismatches 417;
 243 GCTGTGGACCCCAAGACATTTCAATGATTTAGTGAATCATCCAAATCATCAAGGCTTATCCC 302
 1547 GCTAAGAACCCCTGAAGCCAGTATGATGATGATGATGATGATGATGATGATGATGATG 1606
 303 TGTGAGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 362
 1607 AGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1666
 363 CGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 422
 1667 CATGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1726
 423 CTAGTTGAGGTGCTAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 482
 1727 CTTCTGCTCAGCTACAACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1786
 483 CTGGCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 542
 1787 CTGGCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1846
 543 AAACACAGACACTCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 602
 1847 GAACATTTATCTATTTACCACTGATGATGATGATGATGATGATGATGATGATGATGATG 1906
 603 GCTAGTTTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 662
 1907 GCTGAATGATGCTTCCATCTACATGATGATGATGATGATGATGATGATGATGATGATGATG 1966
 663 ATCTATTATGCTGCTATTCAGGCGACCACTGCTGCTTTATGCTGCTTTATGCTGCTTTT 722
 1967 CTACACTATGTTGGCCATTTCCAGGCGACCACTGCTGCTTTGCTGCTTTTCTTACCAAT 2026
 723 CCAGAGCTGCTCAGAAATCAAAATGATTTTGTGCTGCGAGATGATGATGATGATGATGATG 782
 2027 CCCACATGCTGAAATATGAAATGATTTCCATGATTTAGCCCGAGTCGCAAGTGA 2086
 783 CATGCAAAAGCCCGGAGCAAAATTTTGTGCTGCGAGATGATGATGATGATGATGATGATG 842
 2087 CACACCCAGAGCTGTTTAAACAACTTGCACTTATCTCTCACTTCTCTTCAAGATTAT 2146
 843 TTTGGCAAAAGAAATTTCTGATCAGACCAAGATTTCTCAGACAACTTGTATTACCTT 902
 2147 TTTGGTAAACAAATTTCTACCCACAAATTTTGTGAACTTCTTGTGTTGATGATGATG 2206
 903 TGTGCGAGGTGATTTGATGATGATTTGATGATTTGATGATTTGATGATTTGATGATTTGATG 962
 2207 TGCTCTGCTGAGACACTGATGCTTTTGTGAAGATGCTTTGCTGCTTCTTCAAGCT 2266
 963 AACACCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1022
 2267 GACATATAAACTTCACATGATGCTGCTGATGATGATGATGATGATGATGATGATGATGATG 2326
 1023 ACATCTGTGCAAAATATTTCTACATGAGCGAGCGAGTGAATTTCTGCTGAACTCCGGCA 2386
 2327 GCTTCTGTGTTCAAAACATCTCCATGAGGAGCAGGCTATTAAGTCTGGGAAATCCAAGCT 2446
 1083 TTTGACTGGGGAGTGGAGACCAAAATCTGGAATATGCAATGCAATGCAATGCAATGCAATG 1142
 2387 TTTGACTGGGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2446
 1143 TACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1202
 2447 TACAATTTACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2506
 1203 TCAATCCAGAGAGTGAATGCTGCTGCTGAGTGAACCTGATGATGATGATGATGATGATG 1262
 2507 GCTGACCTTCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2566
 1263 AATATTCCTGATGGGCTCAGTGGATTTCACTGCGGCTTTGATGATGATGATGATGATGATG 1322
 2567 GAAATTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2626
 1323 TACATGAAATCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1386
 2627 TACAATGAAATTTATTTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 2686

Db 2327 GCTTCTGTTCAAAACATCTCCACCTGGAGACAGGCTATTAACTCTGGGAAATTCAGGT 2386
 QY 1083 TTTGACTGGGGAGTGAGACCAAAATCTGGAAAAATGCAATCAGCCAACTCTGTAAGG 1142
 Db 2387 TTTGACTGGGGAGCTTCAGTTGAGAACCTTAATGCAATTAATTAATCAGCCCAACCTCCCATC 2446
 QY 1143 TACAGAGTCAGAGATGATGAGGTCCTCAGCAATGTGACAGGAGGTGAGGCTG 1202
 Db 2447 TACAATTTAACGCCATGATGTCCAAATTGCAATGAGTGCTGGCCNAGACCTGTTG 2506
 QY 1203 TCAATTCAGAGACGCTGAAATGCTGCTCTCTGAGGTGACCACTCATCTACCAATAG 1262
 Db 2507 GCTGACCTCTCAGGATGTTGACCTTTTGCTTTCAAACTCTCTAATCTCATTCACCAAG 2566
 QY 1263 AATATTCTGAAATGGGCTCAGCTGGGATTCATCTGGGGTTTGATGCTCTCACCCTATG 1322
 Db 2567 GAAATTCCAATTAATCAATCATCTGGACTTTTCTGAGCAATGATGACCTCAGAGTT 2626
 QY 1323 TACAATGAATATCCATCTGATGACAGGAGGAGA 1359
 Db 2627 TACAATGAATATTTCTTTGATGGCAAAAGACAAA 2663

RESULT 10

US-08-227-108-2
 ; Sequence 2, Application US/08227108
 ; Patent No. 5807726
 ; GENERAL INFORMATION:
 ; APPLICANT: Blanchard, Claire
 ; APPLICANT: Benicourt, Claude
 ; APPLICANT: Junien, Jean-Louis
 ; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA: US/08/227,108
 FILING DATE: 03-APR-1994
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Fanucci, Allan A.
 REGISTRATION NUMBER: 30,256
 REFERENCE/DOCKET NUMBER: 7620-033
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212 790-9090
 TELEFAX: 212 869-8864/9741
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1137 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: Double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..1137
 US-08-227-108-2

Query Match 21.7%; Score 444.2; DB 1; Length 1137;
 Best Local Similarity 62.8%; Pred. No. 3e-117;
 Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

QY 250 ACCCAGAGCATTCATGAATATTTAGTGAATCATCCAAATCAAGGCTATCCCTGTGAGG 309
 Db 26 ACCTGAGTGAACCATGATATTAAGTCAGATGATCACTTCTGGGATACCCAGCTGAGG 85
 QY 310 AATATGAAGTCGAACTGAAGATGGGTATATCTTTCTCTTAAACAGATTCCTCGAGGCC 369
 Db 86 AATATGAAGTTGTGACCGAAGACGGTTATATCTCTGGGATCGACAGATTCCTTATGGGA 145
 QY 370 TAGTGCAACCTAAGAGACAGAGTTCCAGGCTGTGGTGTACTGACAGCATGGCTAGTTG 429
 Db 146 GGAATAATTCAGAGATATATAGCCGGAGACCTTTGCAATTTTGGCAACAGGTTTCTG 205
 QY 430 GAGTGCTAGCAACTGGAATTTCCAACTCTCCCAATAGCTCTGGCTTCATTTCTGGCAG 489
 Db 206 CATCAGCCACAAACTGGATCTCCAACTCTCCCAACAAACAGCTCTGGCTTCATCTGGCCG 265
 QY 490 ATGCTGCTTTTGGAGTGTGGATGGGGAACAGCAGGGAACGCTGCTCGAAGACACA 549
 Db 266 ACGCCGGGTACGACGTGTGGCTGGGGAACAGCAGGGAACACCTGGCCAGGAGGAATC 325
 QY 550 AGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTTCAGTTATGATGAGATGGCTAGGT 609
 Db 326 TGTACTACTCGCCGACTCCGTCGAATTTCTGGGCTTTTCAGCTTTGACGAGATGGCTAAAT 385
 QY 610 TTGACCTTCTGCAAGTGAATACTTTATTTTTCAGAAACGGGCGCAGGAAAGATCTATT 669
 Db 386 ATGACCTTCTCCGCGCACCATTTGACTTCATCTTGAAGAAACGGGACAGGCAAGCTACACT 445
 QY 670 ATGTCGGCTATTTCACAGGGAACCACTGGCTTTTATTTGCAATTTTCCACCATGCCAGAGC 729
 Db 446 ACCTTGGCCATTTCCAGGGAACCACTGGCTTTTCAGCTTTTCCACCAATCCCAGGC 505
 QY 730 TGGCTAGAAATCAAAATGATTTTGTCTTTAGACCATAGCCACTGTTAAGCATGCAAT 789
 Db 506 TGGGAAACGGATCAAAACCTTTCTATGCAATTTAGCTCCCGTTGCCCGGTGAAGTACACCG 565
 QY 790 AAAGCCCGGGAACCAAAATTTTGTGTCGCAGATATGATGATCAAGGGAATGTTGGCA 849
 Db 566 AAACCTGTTAAACAACTCATGCTCGTCCCTTCCTTCAAGCTTATATTGGAA 625
 QY 850 AAAAAGAAATTTCTGATCAGACAGATTTCTCAGACAACTTTGATTTATTTACCTTTGGGCC 909
 Db 626 ACAAATATTTCTACCCACCACTCTTTTGTATCAATTTCTGCGCCACCGAGGTATGCTCCC 685
 QY 910 AGGTGATTTCTGATCAGATTTGATGAATATATCATGTTACTTCTGGGTGGAATCAACACCA 969
 Db 686 GCGAGACGGTGGATCTCTCTGAGCAACGCCCTGTTTATCATTTGTGATTTGACACTA 745
 QY 970 ACAATATGAACATGAGCCGAGCAAGTGTATATGCTGCCACACTCTTGTGCGAACTG 1029
 Db 746 TGAATTTGAACATGAGTGGCTTGGATGTATCTGTCAATAATCCAGCAGGAACATCGG 805
 QY 1030 TGCATAATTTCTACACTGGAGCCAGGAGTGAATTCGTGTAATTCGGGCAATTTGACT 1089
 Db 806 TTCAGAACGCTGCTCCACTGCTGCGAGCAACGCCCTGTTTATCATTTGTGATTTGACT 865
 QY 1090 GGGGAGGTGAGACCAAAATCTGGAATAATGCAATCAGCCAACTCTCTAAGGTACAGAG 1149
 Db 866 GGGGAAGCCAGTTTCAGAACATGATGCACTATCATCAGAGCATGCTCCCTACTACAAC 925
 QY 1150 TCAGAGATGAGCGTCCCTACAGCAATGTGGACAGAGGTGAGGAGGTGAGGCTTTCAATC 1209
 Db 926 TGACAGACATGATGTCCTCAATCGAGTGTGGAAACGGTGGCAACGAGCTTCTGGCGACC 985
 QY 1210 CAGAGACGTTGAAATGCTGCTCTCTGAGGTGACCAACCTCATCTACCAATAAGATATTC 1269
 Db 986 CTCAGATGTTGACCTTTTCTTCCAGCTCCCACTCTCCATCTCATTTACCAAGAGATTC 1045
 QY 1270 CTGAATGGGCTCAGGTGGAATTTTCTGCGGTTTGGATGCTCTCAGCGGTATGCAATG 1329
 Db 1046 CTCCTTCAATCACTTTGAGCTTTATCTGGGCAATGATGCTGCGGCTTCAAGGTTTCAATG 1105

QY 1330 AAATCATCCATCTGATG 1346
 Db 1106 AAATGTTTCCATGATG 1122

RESULT 11
 US-09-073-674-2
 ; Sequence 2, Application US/09073674
 ; Patent No. 5998189
 ; GENERAL INFORMATION:
 ; APPLICANT: Blanchard, Claire
 ; APPLICANT: Benicourt, Claude
 ; APPLICANT: Junien, Jean-Louis
 ; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Warner-Lambert Company
 ; STREET: 2800 Plymouth Road
 ; CITY: Ann Arbor
 ; STATE: Michigan
 ; COUNTRY: U.S.A.
 ; ZIP: 48105

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/073,674
 FILING DATE:
 CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
 NAME: Crissey, Todd M.
 REGISTRATION NUMBER: 37,807
 REFERENCE/DOCKET NUMBER: 5072-D1-66-TWC
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 734 622-7530
 TELEFAX: 734 622-1553
 TELEX:

INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1137 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..1137
 US-09-073-674-2

Query Match 21.7%; Score 444.2; DB 2; Length 1137;
 Best Local Similarity 62.8%; Pred. No. 3e-117;
 Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

QY 250 ACCGAGCAATCATGATATTAGTGAATCATCAACATCAAGCTATCCCTGTGAGG 309
 Db 25 ACCCTGAGTGACCATGATATAGTCAGATGATCACTCTGGGGATACCCAGCTGAGG 85

QY 310 AATATGAGTGGCACTGAGATGGGTATATCCTTTCTGTTAAACAGGATTCCTCGAGGCC 369
 Db 85 AATATGAGTGGTGAACGAGAGCGGTATATCCTTTGGGATCGACAGAAATTCCTTATGGGA 145

QY 370 TAGTGAACCTAAGAGACAGGTTCCAGGCGCTGTGTGTACTGTGAGATGSCCTAGTTG 429
 Db 146 GGNAAATTCAGAGATATAGCCCGAGACCTTGTGATTTTGAACACGGTTTGTCTG 205

QY 430 GAGGTGCTAGCACTGATTTTCAACCTGCCCAACATAGCTGGGCTTCATTTCTGGCAG 489
 Db 206 CATCAGCCCAAACTGGATCTCAACCTGCCCAACACAGCTGGCTTCATCTCTGGCG 265

QY 490 ATGCTGTTTACCTGTGGTGGGACAGAGGAGAACCCCTGCTCGAAGAACACA 549

Db 266 ACGCCGGGTACGACGTGTGCTGGGGAACAGCAGGGGCAACACCTGGGCCAGGGAATC 325
 QY 550 AGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTTCAGTTATGATGAGATGGCTAGGT 609
 Db 326 TGTACTACTCGCCGACTCGTGAATTCCTGGGCTTTTCAGCTTTGACGAGATGGCTAAAT 385

QY 610 TTGACCTTCTGCACTGATGAATTTATTTTTCAGAGAAACGGGCCAGGAAAGATCTATT 669
 Db 386 ATGACCTTCCCGCCACCAATTTGACTTTCATCTTGAAGAAACGGGACAGGCAAGCTACCT 445

QY 670 ATGTGCGCTATTACAGGGGCAACCAATGGGCTTTATTTGCAATTTTCCACCATCCAGAGC 729
 Db 446 ACGTTGGCAATTCAGGGGCAACCAATGGGCTTTTCATCGCTTTTCCACCAATCCCAAGC 505

QY 730 TGGCTCAGAAATCAAAATGATATTTTGTCTTTAGCACCCATAGCCATGTTTAAAGCATGCAA 789
 Db 506 TGGCGAAACGGATCAAAACCTTCTATGCAATAGTCTCCGTTGCCCGTGAAGTACACCG 565

QY 790 AAAGCCCGGGAGCAAAATTTTGTCTGCGCAGATATGATCAAGGGATGCTTTTGGCA 849
 Db 566 AAACCTGTTAAACAACTCATGCTGCTCCCTTCCTTCAAGCTTATATTTGGAA 625

QY 850 AAAAAGAAATTCGTATCAGACCAAGATTTCTCAGCAACTGTGTTATTTACCTTTTGGCC 909
 Db 626 ACAAATATTTCTACCCACCACTTCTTTGATCAATTTCTCGCCACCGAGGTATGCTCCC 685

QY 910 AGGTGATTTCTGATCAGATTTGTTAGTAATATATCTTCTCTGCGTTCCTTCAAGCTTAT 969
 Db 686 GCGAGCGGTGGATCTCTCTGAGCAACGCTCTGTTATCATTTGTGATTTGACACTA 745

QY 970 ACAATATGAACATGAGCCGAGCAAGTGTATATGCTGCCCACTCTTTGCTGGAACATCTG 1029
 Db 746 TGAACCTTGAACATGAGTGGCTTGGATGTGTATCTGTCAATAATCCAGCAGGAACATCGG 805

QY 1030 TGCAAAATATTTCTACCTGAGCCAGCAGTGAATTTCTGGTGAATCTCCGGGCAATTTGACT 1089
 Db 806 TTCAAGACGTGCTCCCACTGGTCTCCCGCTGTTAAGTCTGGGAAGTTCCAGCTTTTACT 865

QY 1090 GGGGAGTGAGACCAAAATCTCGAAAAATGCAATCAGCCAACTCCTGTAAGGTACAGAG 1149
 Db 866 GGGGAGCCCGAGTTCAAGACATGATGCACTATCATCAGAGCATGCTCCTCTACTACAAC 925

QY 1150 TCAGATATGACGGTCCCTACAGCAATGTGACAGAGGTGAGGAGGTGAGGCTTCTTCAATC 1209
 Db 926 TGACAGACATGATGTCGAATCGCAATCGCAGTGTGGAAACGGTGGCAACGACTTCTGCGCCGCC 985

QY 1210 CAGAGACGTGAAAAATGCTGCTCTCTGAGGTGACCAACCTCATCTACCAATAAGAAATTC 1269
 Db 986 CTCAGATGTTGACCTTTTGTCTTCCAGCTCCCAATCTCATTTACCAAGAGATTC 1045

QY 1270 CTGAATGGGCTCAGGTGATTTTCATCTGGGGTTTGGATGCTCTCCCGTATGTACAATG 1329
 Db 1046 CTCCTTACATCACTTTGGACTTTATCTGGGCCATGATGATGCCCTCAAGCGGTTTACAATG 1105

QY 1330 AAATCATCCATCTGATG 1346
 Db 1106 AAATGTTTCCATGATG 1122

RESULT 12
 US-08-227-108-4
 ; Sequence 4, Application US/08227108
 ; Patent No. 5807726
 ; GENERAL INFORMATION:
 ; APPLICANT: Blanchard, Claire
 ; APPLICANT: Benicourt, Claude
 ; APPLICANT: Junien, Jean-Louis
 ; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York

```

; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/227,108
; FILING DATE: 03-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fanucci, Allan A.
; REGISTRATION NUMBER: 30,256
; REFERENCE/DOCKET NUMBER: 7620-033
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1140 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1140
; US-08-227-108-4

Query Match 21.7%; Score 444.2; DB 1; Length 1140;
Best Local Similarity 62.8%; Pred. No. 3e-117;
Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

QY 250 ACCCAGAGCATTCATGAATATTAGTGAATCATCCAAATCAAGGCTATCCCTGTGAGG 309
DB 29 ACCCTGAAGTACCATGATATAGTCAGATGATCACTACTGGGATACCCAGCTGAGG 88
QY 310 AATATGAAGTGCACAACTGAAGATGGGTATATCTTTCTGTAAACAGGATTCCTCGAGGCC 369
DB 89 AATATGAAGTGTGACCGAAGACGCTTATATCTTGGGATCGACAGAAATCTCTATGGGA 148
QY 370 TAGTGCACCTAAGAGACAGAGTTCAGGCGCTGTGTTACTGCACATGGGCTAGTTG 429
DB 149 GGAAAAATTCAGAAATATAGCGCGAGACCTGTGTGATTTTGCAACAGGTTGCTCG 208
QY 430 GAGGTGCTAGCAACTGGATTTCCAACTGCCCCCAACATAGCTGGGCTTCATTTCTGGCAG 489
DB 209 CATCAGCCACAACTGGATCTCCAACTGCCCCCAACAGCCTGGGCTTCATCTGGCCG 268
QY 490 ATGCTGGTTTTCAGTGTGGATGGGACAGCAGGGGAACGCTGTCTCGAAACACA 549
DB 269 ACGCGGGGTACGAGTGTGGCTGGGGAACAGCAGGGGAACACCTGGGCGAGGGAATC 328
QY 550 AGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTTCAGTTATGATGAGATGGCTAGGT 609
DB 329 TGTACTACTCGCCGACCTCCGTGAAATCTGGGCTTTTCAGTTTTCAGGATGGCTAAAT 388
QY 610 TTGACCTTCCTGAGTGAATACTTTATTTTGAGAAAAACGGGCGCAGAAAAAGATCTATT 669
DB 389 ATGACCTTCCGCGCCACCAATTGACTTCATCTTTGAAGAAAAACGGGACAGGACAAGTACACT 448
QY 670 ATGTCGGGTATTTCAGGGGACCAACCATGGGCTTTTATTTGATTTTTCACCATGCGAGC 729
DB 449 ACGTTGGCCATTCACGGGCGACCACTATGTTTCATCGCTTTTCACCAATCCCAAGC 508
QY 730 TGGCTCAGAAATCAAAATGATTTTGTCTTTAGCACCACATAGCCACTGTTAAGCATGCAA 789
DB 509 TGGCGAAACGGATCAAAACCTTCTATGATGATAGTCCCGTTGCCACCGTGAAGTACACCG 568
QY 790 AAAGCCCCGGGACCAAAATTTTGTGCTGCCAGATATGATGATCAAGGGATGTTTGCA 849

; 569 AAACCCCTGTTAAACAACTCATGCTCGTCCCTTCCTTCAGCTTATATTTGGAA 628
; 850 AAAAAGAAATTTCTGTATCAGACCAAGATTTCTCAGACAACTTGTATTATTACCTTTGTGGCC 909
; 629 ACAAAATATTTCTACCCACACCACTTCTTTGATCAATTTCTGCCACCGAGGTATGCTCCC 688
; 910 AGGTGATTTCTGATCAGATTTGTAGTAATATCATGTTTACTTCTGGGTGGATTCAACACCA 969
; 689 GCGAGACGGTGGATCTCTCTGCGACCAAGCCCTGTTTATCATTTTGGATTGCACTA 748
; 970 ACAATATCAACATGAGCGCAGCAAGTGTATATGCTGCCCACTCTTTGCTGGAAACATCTG 1029
; 749 TGAACCTTCAACATGAGTGGCTTGGATGTGTATCTGTACATAATATCCAGCAGGAACATCGG 808
; 1030 TGCAAAATATTTCTACACTGAGCGCAGGAGTGAATTTCTGGTGAACCTCCGGSCATTTGACT 1089
; 809 TTCAGACGTGTCTCCACTGGTCCAGAGCTGTAGTCTGGGAAGTTTCAAGCTTTTGAAT 868
; 1090 GGGGAGTGCAGACCAAAATCTGAAAAATCAATCAGCCAACTCTGTAAAGGTACAGAG 1149
; 869 GGGGAAGCCAGTTTCAGAAACATGATGCACTATCATCAGAGCATGCTCCCTTACTACAACC 928
; 1150 TCAGAGATATGACGCTCCCTACAGCAATGTGCGAGGAGGTTCAGGATGCGCTTTCAATC 1209
; 929 TGACAGACATGATGTCGAATCGCAGTGTGGAACGGTGGCAAGCTTTGCTGCCGACC 988
; 1210 CAGAAAGACGTGAAAAATGCTGCTCTGAGGTGCAACCACTCATCTACCATAGAATATTC 1269
; 989 CTCAGATGTTGACCTTTTGTCTTCCAAAGCTCCCAATCTCATTTACCAAGGAATTC 1048
; 1270 CTGAATGGCTCAGCTGGATTTTCATCTGGGCTTTGGATGCTCTCCACCGTATGTACATG 1329
; 1049 CTCCTTCAACATCACTTGGACTTTATCTGGGCCATGAGTGGCCCTCAAGCGGTTTACATG 1108
; 1330 AAATCATCATCTGATG 1346
; 1109 AAATGTTTCCATGATG 1125

RESULT 13
US-09-073-674-4
; Sequence 4, Application US/09073674
; Patent No. 5998189
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Warner-Lambert Company
; STREET: 2800 Plymouth Road
; CITY: Ann Arbor
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 48105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,674
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Crissey, Todd M.
; REGISTRATION NUMBER: 37,807
; REFERENCE/DOCKET NUMBER: 5072-D1-66-TWC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 734 622-7530
; TELEFAX: 734 622-1553

```

```
TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1140 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1140
; US-09-073-674-4

Query Match      21.7%; Score 444.2; DB 2; Length 1140;
Best Local Similarity 62.8%; Pred. No. 3e-117;
Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

QY 250 ACCCAGAGCATTCATGATATAGTGAATCATCAACATCAAGCTATCCCTGTGAGG 309
DB 29 ACCCTGAAGTGACCATGAATATAGTCAGATGATACCTACTGGGGATACCCAGCTGAGG 88
QY 310 AATATGAAGTCGCAACTGAAGATGGGTATATCCCTTTCTGTAAACAGGATTCCTCGAGGCC 369
DB 89 AATATGAAGTTGTGACGAGACGGTTATATCCTTGGGATCGACAGAAATTCCTTATGGGA 148
QY 370 TAGTGCACACTAAGACAGAGGTTCCAGGCTGTGTGTACTGACGATGSCCTAGTTG 429
DB 149 GGAATAATTGAGAAATATAGCCGAGACCTGTTCATTTTTCACACAGGTTTGTCTCG 208
QY 430 GAGGTGCTAGCACTGGATTTCCAACTGCGCCCAACATAGCTGGGCTTCATTCTGGCAG 489
DB 209 CATCAGCCCAAACTGGATCTCCAACTGCGCCCAACAGAGCTGGCTTCATCTGGCG 268
QY 490 ATGCTGGTTTGTACGTTGGTGGGAAACAGCAGGGGAAACGCTGTCTCGAAACACA 549
DB 269 ACCCGGGTACAGCTGTGGCTGGGAAACAGCAGGGGCAACACCTGGCGCCAGGGAATC 328
QY 550 AGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTTCAGTTATGATGATGAGTGTAGT 609
DB 329 TGTACTACTCGCCGACTCGCTCGAATTTCTGGGCTTTTCAGTTTACGAGATGGCTAAT 388
QY 610 TTGACCTTCTGAGTGATATACTTTATTTTTCAGAAACGGGCGACGAAAGATCTATT 669
DB 389 ATGACCTTCCGCGCCACCATTTGACTTTCATCTTGAAGAAACGGGACAGGACGACTACAT 448
QY 670 ATGTCGCTATTTCACAGGCGCACCATGGCTTTTATTCATTTTCCACCATGCCAGAGC 729
DB 449 ACCTTGGCCATTCGAGGCGACCACTATGGTTTCATCGCTTTTCCACCATCCAGC 508
QY 730 TGGCTCAGAAATCAAAATGATTTTGTCTTACGACCCATAGCCACTGTTAAGCATGCAA 789
DB 509 TGGCGAAACGGATCAAAACCTTCTATGCAATAGTCTCCGTTGCGACCGTGAAGTACCG 568
QY 790 AAAGCCCGGACCAAAATTTTGTCTGTCGAGATGATGATCAAGGATTTGTTGCA 849
DB 569 AAACCTGTTTAAACAACTCATGCTGCTGCTGCTGCTTCTTCAAGCTTATTTTGA 628
QY 850 AAAAAGAAATTTGTATCAGACAGCAATTTCTCAGACAACTGTTATTTACCTTTGTGGCC 909
DB 629 ACAAATATTTCTACCCACCACTTCTTTGATCAATTTCTCGCCACCGAGGTATGCTCC 688
QY 910 AGGTGATTTCTGATCAGATTTGTAGTAAATATCATGTTACTTCTGGGTGATTCACACCA 969
DB 689 GCGAGACGGTGGATTCCTCTCGAGCAACGCTGTTTATCATTTTGTGATTTGACATA 748
QY 970 ACAATATGAACATGAGCCGAGCAAGTGTATGCTGCCACACTCTTGTGGAACATCTG 1029
DB 749 TGAACTTGAACATGAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 808
QY 1030 TGAATAATTTCTACACTGGAGCGAGCTGATGATGATGATGATGATGATGATGATGATGAT 1089
DB 809 TTCAAGACGCTGCTCCACTGCTGCCAGCTGTTAAGTCTGGGAAGTTCACAGCTTTTGACT 868

1090 GGGGAGTGAGACCAAAATCTGAAAATTCGAATCAATCAGCCACTCCTGTGAAGTACAGAG 1149
DB 869 GGGGAGGCCAGTTTCAGAAACATGATGCACTATCATCAGACATGCTTCCCTACTACACC 928
QY 1150 TCAGAGATATGACGGTCCCTACAGCAATGTGAGACAGGAGTCAAGGACTGGCTTTCAAATC 1209
DB 929 TGACAGACATGATGTGCCAATCCAGTGTGGAACGGTGGCAACGACTTCTGCTGGCGACC 988
QY 1210 CAGAAGACGTGAAAATCTCTCTGAGGTGAGCAACCTCATCTACCATCAAGATATTC 1269
DB 989 CTACAGATGTGACCTTTTCTTCCAGCTCCCAATCTCATTTTACCAAGAGATTC 1048
QY 1270 CTGAATGGCTCAGCTGATTTTCATCTGGGTTTGGATGCTCTCCAGCTGATGACATG 1329
DB 1049 CTCCTTCAACATCACTTGGACTTTTATCTGGGCCATGATGCCCTCAAGCGTTTACAATG 1108
QY 1330 AATCATCCATCTGATG 1346
DB 1109 AATGTTTCCATGATG 1125

RESULT 14
US-08-227-108-6
; Sequence 6, Application US/08227108
; Patent No. 5807726
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/227,108
; FILING DATE: 03-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Panucci, Allan A.
; REGISTRATION NUMBER: 30,256
; REFERENCE/DOCKET NUMBER: 7620-033
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1146 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; US-08-227-108-6

Query Match      21.7%; Score 444.2; DB 1; Length 1146;
Best Local Similarity 62.8%; Pred. No. 3e-117;
Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

QY 250 ACCCAGAGCATTCATGATATAGTGAATCATCAACATCAAGGCTATCCCTGTGAGG 309
DB 26 ACCCTGAAGTGACCATGAATATAGTCAGATGATACCTACTGGGGATACCCAGCTGAGG 85
QY 310 AATATGAAGTCGCAACTGAAGATGGGTATATCCCTTTCTGTAAACAGGATTCCTCGAGGCC 369
```

Db 86 AATATGAAGTGTGACCGAGACGGTTATATCTTGGATCGACAGAAATCTTATGGGA 145
 Qy 370 TAGTGCAACCTAAGAGACAGGTTCAGGCTGTGGTGTGTTACTGACAGATGGCTAGTTG 429
 Db 146 GGAATAATTCAGAGAAATATAGCGCGGAGACCTGTGTCATTTTGGCAACGGTTGTCTG 205
 Qy 430 GAGGTGTAGCAACTGGATTTCACCACTGCCCAATAGCTGGGCTTCATTTCTGCGAG 489
 Db 206 CATCAGCCCAAACTGGATCTCCAACTGCGCAACACAGCCTGGCTTCATCTGCGCG 265
 Qy 490 ATGCTGTTTGAAGTGTGAGATGGGAAACAGCAGGGGAAACGGCTGTCTCGAAACACA 549
 Db 266 ACGCGGGTACAGCTGTGGCTGGGAAACAGCAGGGGAAACACCTGGCGCAGGAGAAATC 325
 Qy 550 AGACACTCTCCATAGACCAAGATGTTCTGGGCTTTTTCAGTTATGATGAGATGGTAGGT 609
 Db 326 TGTACTACTCCCGCAGCTCCCTCGAATCTGGGCTTTTTCAGTTTGCAGATGGCTAAT 385
 Qy 610 TTGACCTTCTGCAAGTATAACTTTATTTTGCAGAAACCGGCGCAGGAAAGATCTATT 669
 Db 386 ATGACCTTCCCGCAGCTCCCTCGAATCTGGGCTTTTTCAGTTTGCAGATGGCTAAT 385
 Qy 670 ATGCTGGCTATTCAGCGGACCAACCTGCTTATTTGATTTTCCACCATGCCAGC 729
 Db 446 ACGTTGGCCATTCAGGGCACCACCTGTTTTCATGCTGCTTTCACCAATCCAGC 505
 Qy 730 TGGCTCAGAAATCAAAATGTTATTTGCTTTAGCACCATAGCCACTGTTAAGCATGCAA 789
 Db 506 TGGGAAACCGATCAAAACCTTCTATGATAGTCCGTTGCCACCGTGAAGTACACCG 565
 Qy 790 AAAGCCCGGACCAAAATTTTGTGCTGCCAGATATGATCATCAAGGATTTTGGCA 849
 Db 566 AAACCTGTTAAACATCATGCTGCTGCTTCTGCTTCTTCAAGCTTATTTGGAA 625
 Qy 850 AAAAGAAATTTCTGTATCAGACCAAGATTTCTCAGACAACTTTGTTATTTACCTTTGGCC 909
 Db 626 ACAAAATATTCACCCACACCTTCTTTGATCAATTTCTGCCACCGAGTATGCTCCC 685
 Qy 910 AGGTGATTTGATCAGATTTGTAGTAAATATCATGTTACTTCTGGTGGATTCACACCA 969
 Db 686 GCGAGACGGTGGATCTCTCTGACGAAACGGCTGTTTATCATTTGTTGATTTGACACTA 745
 Qy 970 ACAATATGAATGAGCGGACGCAAGTGTATGCTGCCACACTCTTCTGTTGAACTCTG 1029
 Db 746 TGAATTTGAACATGATGCTGCTGGATGTATCTGTACATAATCCAGCAGGACATCGG 805
 Qy 1030 TGCAAAATATTTCACTGAGGCGGAGGAGTGAATTTCTGGTGAATCTGGGCACTTGAAT 1089
 Db 806 TTCAGACGTGCTCCACTGGTCCAGGCTGTTAAAGTCTGGGAAATTTCAAGCTTTTGAAT 865
 Qy 1090 GGGGAGTGAGACCAAAATCTGGAATAATGCAATTCAGCCAACTCTCTGTAAGGTACAGAG 1149
 Db 866 GGGGAGGCCAGTTCCAGAACATGATGCACTATCATCAGAGATGCTCCCTACTACACC 925
 Qy 1150 TCAGAGATATAGCGTCCCTAAGCAATGTGACAGGAGGTGACAGCTGGCTTTCAAATC 1209
 Db 926 TCACAGACATGATGTGCAAAATGCGAGTGTGGAACGGTGGCAACGACTTGTGCGCCGACC 985
 Qy 1210 CAGAAAGCTGAAATGCTGCTCTGAGGTGACCAACCTCATCTACCATAGAAATATTC 1269
 Db 986 CTCACGATGTTGACCTTTGCTTTTCCAAAGCTCCCAATCTCATTTACCAAGGAGATTC 1045
 Qy 1270 CTGAATGGGCTCACGTGATTTTCACTGCGGTTTGGATGCTCTCACCGTATGTACAAATG 1329
 Db 1046 CTCCTTACAAATCACTTGGACTTTATCTGGGCAATGATGCTCCCTCAAGCGGTTTACAATG 1105
 Qy 1330 AATCATCATCTGATG 1346
 Db 1106 AATTTGTTTCCATGATG 1122

RESULT 15

US-09-073-674-6
 ; Sequence 6, Application US/09073674
 ; Patent No. 5998189
 ; GENERAL INFORMATION:
 ; APPLICANT: Blanchard, Claire
 ; APPLICANT: Benicourt, Claude
 ; APPLICANT: Junien, Jean-Louis
 ; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Warner-Lambert Company
 ; STREET: 2800 Plymouth Road
 ; CITY: Ann Arbor
 ; STATE: Michigan
 ; COUNTRY: U.S.A.
 ; ZIP: 48105
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/073,674
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Crissey, Todd M.
 ; REGISTRATION NUMBER: 37,807
 ; REFERENCE/DOCKET NUMBER: 5072-D1-66-TWC
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 734 622-7530
 ; TELEFAX: 734 622-1553
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1146 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA to mRNA
 ; US-09-073-674-6

Query Match 21.7%; Score 444.2; DB 2; Length 1146;
 Best Local Similarity 62.8%; Pred. No. 3e-117;
 Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;
 Qy 250 ACCGAGACCATTCATGATATTTAGTGAATCATCAACATCAAGCTATCCCTGTGAGG 309
 Db 26 ACCCTGAAGTGAACCAATGATATAGTCAAGTATCACTCTGGGATACCACGTGAGG 85
 Qy 310 AATATGAAGTTCGCAACTGGAATGATGTTTCTTCTGTTTAAACAGGATTTCTCGAGGCC 369
 Db 86 AATATGAAGTTCGACCGAAGACGTTATATCTTTGGGATCGACAGAATTCCTTATGGGA 145
 Qy 370 TAGTGCAACTAAGAGACAGGTTCAGGCTGTGGTGTGTTACTGACAGATGGCTAGTTG 429
 Db 146 GGAATAATTCAGAGAAATATAGCGCGGAGACCTGTGTCATTTTGGCAACGGTTGTCTG 205
 Qy 430 GAGGTGTAGCAACTGGATTTCACCACTGCCCAACATAGCTGGGCTTCATTTCTGGCAG 489
 Db 206 CATCAGCCCAAACTGGATCTCCAACTGCGCAACACACAGCTGGCTTCATCTGCGCG 265
 Qy 490 ATGCTGTTTGAAGTGTGAGATGGGAAACAGCAGGGGAAACGGCTGTCTCGAAACACA 549
 Db 266 ACGCGGGTACAGCTGTGGCTGGGAAACAGCAGGGGAAACACCTGGCGCAGGAGAAATC 325
 Qy 550 AGACACTCTCCATAGACCAAGATGTTCTGGGCTTTTTCAGTTATGATGAGATGGTAGGT 609
 Db 326 TGTACTACTCGCCGACTCCGTCGAATCTGGGCTTTTTCAGTTTGCAGATGGCTAAT 385
 Qy 610 TTGACCTTCTGCAAGTATAACTTTATTTTGCAGAAACCGGCGCAGGAAAGATCTATT 669
 Db 386 ATGACCTTCCCGCAGCTCCCTCGAATCTGGGCTTTTTCAGTTTGCAGATGGCTAAT 445

QY 670 ATGTGGCTATTACAGGCGACCAACATGGGCTTTATTGGCATTTTCCACCATGCCAGGC 729
Db |||||
QY 446 ACGTTGGCCATTCACGGGACCAACCAATGGTTTCATCGGCTTTCCACCAATCCCAAGC 505
Db |||||
QY 730 TGGCTCAGAAATCAGAAATGATTTTCTTTAGCACCCATAGCCACTGTTAAGCATGCAA 789
Db |||||
QY 506 TGGGAAACGGATCAAAACCTTCTATGCAATAGCTCCGTTGCCCGTGAAGTACACCG 565
QY 790 AAAGCCCGGACCAAAATTTTGGTGTGCGCAGATATGATGATCAAGGATGTTTGGCA 849
Db |||||
QY 566 AAACCTGTGTAACAAACTGATGCTGCTCCCTTCCTTCAAGCTTATATTGGAA 625
QY 850 AAAAGAAATTTCTGTATCAGACCAAGATTTCTCAGACAACTGTTATTATTACCTTTGTGGCC 909
Db |||||
QY 626 ACAAAATATTCTACCCACACCACTTCTTTGATCAATTTCTGGCCACCGAGGTATGCTCCC 685
QY 910 AGGTGATTTCTGTATCAGATTTGTAGTAATATCATGTTTACTTCTGGTGGATTCACACCA 969
Db |||||
QY 686 GCGAGCGGTGGATCTCCTCTGACGAAACGCGCTGTTTATCATTTGGGATTTGACACTA 745
QY 970 ACAATATGAACATGAGCGGAGCAAGTGTATATGTGCCACACTCTTGTGGAACTCTG 1029
Db |||||
QY 746 TGAATTTGAACATGAGTGGCTTGGATGTGTATCTGTACATATCCAGCAGGAACATCG 805
QY 1030 TGCAGAAATTTCTACACTGAGCGGACGAGTGAATTTCTGTGAACTCCGGGCAATTTGACT 1089
Db |||||
QY 806 TTCAGAACGTCCTCCACTGTGTCAGGCTGTTAAGTCTGGAAAGTTCCAAAGCTTTTACT 865
QY 1090 GGGGAGTGAAGACCAAAATCTGGAAATGCAATCAGCCAACTCTCTGTAAGGTACAGAG 1149
Db |||||
QY 866 GGGGAGCCGAGTTCAGAACATGATGCACTATCATCAGACATGCTCCCTACTACAAC 925
QY 1150 TCAGAGATATGACGGTCCCTACAGCAATGTGGACAGAGGTGAGGCTGAGGCTGCTTCAAATC 1209
Db |||||
QY 926 TGACAGACATGCAATGTGCCAATCGCAGTGTGGAAACGTTGGCAACGACTTGTGCGCGAC 985
QY 1210 CAGAGACGTAAGAAATGCTGCTCTCAGGTGACCACTCTATCCTACCAATAGAAATTC 1269
Db |||||
QY 986 CTCAGATGTGACCTTTGCTTCCAAAGCTCCCAATCTCATTTACCAAGGAATTC 1045
QY 1270 CTGAATGGGCTCAGTGGATTTCAATCTGGGTTTGGATGCTCTCAACCGTATGTACAATG 1329
Db |||||
QY 1046 CTCCTTACAATCACTTGGACTTTATCTGGGCCATGGATGCCCTCAAGCGGTTTACAATG 1105
QY 1330 AATCATCCATCTGATG 1346
Db |||||
QY 1106 AATTGTTCCATGATG 1122

Search completed: February 19, 2004, 10:33:53
Job time : 135.478 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 08:29:29 ; Search time 739.122 Seconds
(without alignments)
10186.874 Million cell updates/sec

Title: US-10-042-431-45
Perfect score: 2044
Sequence: 1 gtcgaccacgctccggg.....aaaaaaaggcgccgc 2044

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2449703 seqs, 1841816367 residues
Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq.*
17: /cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq.*
18: /cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2044	100.0	2044	11	US-09-759-130B-415
2	2044	100.0	2044	14	US-10-042-431-45
3	1993.8	97.5	2035	10	US-09-835-996A-20
4	1343.4	65.7	1384	10	US-09-835-996A-9
5	1269	62.1	1269	11	US-09-759-130B-416
6	1269	62.1	1269	14	US-10-042-431-46
7	1126	55.1	1126	10	US-09-811-825-1
8	890.4	43.6	18554	10	US-09-811-825-3
9	527.8	25.8	2481	10	US-09-880-107-3878
10	527.8	25.8	2858	15	US-10-198-846-13431
11	527.8	25.8	6952	13	US-10-133-013-144
12	525.4	25.7	2853	13	US-09-971-392-116
13	462.2	22.6	2806	12	US-10-108-260A-1262
14	456.8	22.3	1411	12	US-10-420-564-3
15	454.2	22.2	1194	13	US-10-312-088-7

16	453.6	22.2	1134	12	US-10-420-564-4
17	450.8	22.1	1194	14	US-10-056-744B-3
18	450.8	22.1	1325	14	US-10-056-744B-1
19	450.4	22.0	1134	12	US-10-420-564-1
20	449.8	22.0	8324	12	US-10-420-564-5
21	443.6	21.7	1200	13	US-10-221-097-9
22	424.8	20.8	1260	13	US-10-015-115-31
23	414.2	20.3	1360	14	US-10-003-302-1
24	346.4	16.9	1267	12	US-10-051-874-27
25	344.8	16.9	1267	12	US-10-051-874-29
26	332.8	16.3	1098	13	US-10-312-088-6
27	327.8	16.0	1195	12	US-10-051-874-31
28	303.6	14.9	634	15	US-10-198-846-9462
29	300.4	14.7	1377	13	US-10-015-115-29
30	196.6	9.6	495	15	US-10-060-036-4296
31	192.8	9.4	839	15	US-10-198-846-6768
32	171.2	8.4	534	15	US-10-060-036-4334
33	165.4	8.1	510	15	US-10-060-036-4326
34	165.4	8.1	522	15	US-10-060-036-4325
35	161.8	7.9	635	13	US-10-099-926-1834
36	161.8	7.9	635	14	US-10-033-528-1834
37	152.6	7.5	460	12	US-10-062-674-253
38	128.4	6.3	362	10	US-09-960-352-13032
39	124.4	6.1	624	13	US-10-027-632-214253
40	124.4	6.1	624	14	US-10-027-632-214253
41	122.6	6.0	419	11	US-09-918-995-4162
42	113	5.5	1242	12	US-10-260-238-704
43	104.2	5.1	22067	14	US-10-003-302-3
44	100.4	4.9	371	11	US-09-918-995-37615
45	89.4	4.4	367	10	US-09-960-352-13508

ALIGNMENTS

RESULT 1

US-09-759-130B-415
Sequence 415, Application US/09759130B
Publication No. US2003002279A1
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc.
APPLICANT: McCarthy, Sean A
APPLICANT: Fraser, Christopher C
APPLICANT: Sharp, John D
APPLICANT: Barnes, Thomas S
APPLICANT: Kirst, Susan J
APPLICANT: Mackay, Charles R
APPLICANT: Myers, Paul S
APPLICANT: Leiby, Kevin R
APPLICANT: Wrighton, Nicolas
APPLICANT: Goodearl, Andrew
APPLICANT: Holtzman, Douglas A
TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
FILE REFERENCE: MPI00-5350NMIM
CURRENT APPLICATION NUMBER: US/09/759,130B
CURRENT FILING DATE: 2002-09-16
PRIOR FILING DATE: US 09/479,249
PRIOR FILING DATE: 2000-01-07
PRIOR FILING DATE: US 09/559,497
PRIOR FILING DATE: 2000-04-27
PRIOR FILING DATE: US 09/578,063
PRIOR FILING DATE: 2000-05-24
PRIOR FILING DATE: US 09/333,159
PRIOR FILING DATE: 1999-06-14
PRIOR FILING DATE: US 09/596,194
PRIOR FILING DATE: 2000-07-14
PRIOR FILING DATE: US 09/342,364
PRIOR FILING DATE: 1999-06-29
PRIOR FILING DATE: US 09/608,452
PRIOR FILING DATE: 2000-06-30
PRIOR FILING DATE: US 09/393,996

```

; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/502,871
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/420,707
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 415
; LENGTH: 2044
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-759-130B-415

Query March
Best Local Similarity 100.0%; Score 2044; DB 11; Length 2044;
Matches 2044; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCGACCCAGCGCTCGGGGAATTCGAGCAGGGAATATGTAAGAGTTTAAACCCAC 60
DB 1 GTCGACCCAGCGCTCGGGGAATTCGAGCAGGGAATATGTAAGAGTTTAAACCCAC 60
QY 61 AAATTCCTCTTACTTTAGAAATAGTTGTTACATTCGAGGAAATTAATGAGATGTT 120
DB 61 AAATTCCTCTTACTTTAGAAATAGTTGTTACATTCGAGGAAATTAATGAGATGTT 120
QY 121 GGACCATGTTGGAAACCTCTCAAGACAGTGGATGTTCTCACACAGAAATGGAATGGC 180
DB 121 GGACCATGTTGGAAACCTCTCAAGACAGTGGATGTTCTCACACAGAAATGGAATGGC 180
QY 181 TTCTGATTTCTGTTGGCGTATGTTTCAGAGAAATGTGAATTCAGTACATATGCCAATTA 240
DB 181 TTCTGATTTCTGTTGGCGTATGTTTCAGAGAAATGTGAATTCAGTACATATGCCAATTA 240
QY 241 AAGCTGTGGACCCGAGAGCTTCAAGATGATGTAATGTAATGTAATGTAATGTAATGTA 300
DB 241 AAGCTGTGGACCCGAGAGCTTCAAGATGATGTAATGTAATGTAATGTAATGTAATGTA 300
QY 301 CCTGTGAGGAATATGAAGTCCAACTCAAGATGGGTATATCTCTGTTTAAAGGATTC 360
DB 301 CCTGTGAGGAATATGAAGTCCAACTCAAGATGGGTATATCTCTGTTTAAAGGATTC 360
QY 361 CTCGAGGCTAGTGCAACCTTAAGAGACAGGTTCCAGGCTGTGGTGTACTGCAGCATG 420
DB 361 CTCGAGGCTAGTGCAACCTTAAGAGACAGGTTCCAGGCTGTGGTGTACTGCAGCATG 420
QY 421 GCCTAGTTGAGGCTAGCACTGCAATTCGATTTCCAACTGCGCCCAATAGCCTGGCTTCA 480
DB 421 GCCTAGTTGAGGCTAGCACTGCAATTCGATTTCCAACTGCGCCCAATAGCCTGGCTTCA 480
QY 481 TTCTGGCAGATGCTGGTTTTCAGCTGGATGGGAAACAGCAGGGGAAACGCTGGCTTC 540
DB 481 TTCTGGCAGATGCTGGTTTTCAGCTGGATGGGAAACAGCAGGGGAAACGCTGGCTTC 540
QY 541 GAAACACAGACACTCTCCATAGACCAAGATGATGTTCTGGGCTTTTCAAGTATGATGAGA 600
DB 541 GAAACACAGACACTCTCCATAGACCAAGATGATGTTCTGGGCTTTTCAAGTATGATGAGA 600
QY 601 TGGCTAGTTTACCTTCTGCTGATGATTAATTTTCAGAGAAACGGGCGAGGAAA 660
DB 601 TGGCTAGTTTACCTTCTGCTGATGATTAATTTTCAGAGAAACGGGCGAGGAAA 660
QY 661 AGATCTATTATGCTGGCTATTTCAGGCGACCACTGATGCTTATGCTATGCTTTCACCA 720
DB 661 AGATCTATTATGCTGGCTATTTCAGGCGACCACTGATGCTTATGCTATGCTTTCACCA 720
QY 721 TGCCAGAGCTGGCTCAGAAATCAAAATGATTTTGTGCTTACACCCATAGCCACTGTTA 780
DB 721 TGCCAGAGCTGGCTCAGAAATCAAAATGATTTTGTGCTTACACCCATAGCCACTGTTA 780
QY 781 AGCATGCAAAAGCCCGGAGCCAAATTTTGTGCTGCGCAGATATGATCAAGGAT 840
DB 781 AGCATGCAAAAGCCCGGAGCCAAATTTTGTGCTGCGCAGATATGATCAAGGAT 840

```

```

841 TGTGTGGCAAAAGAAATTTCTGTATCAGACAGATTTCTCAGACAACTTGTATTATACC 900
841 TGTGTGGCAAAAGAAATTTCTGTATCAGACAGATTTCTCAGACAACTTGTATTATACC 900
901 TTTGTGGCCAGGGATTTCTGTATCAGATTTGTAGTAAATATATCTCTGGGTGGAT 960
901 TTTGTGGCCAGGGATTTCTGTATCAGATTTGTAGTAAATATATCTCTGGGTGGAT 960
961 TCAACACCAACATATGAAATGAGCGGAGCAAGTGTATATGTCGCCACACTCTTGCTG 1020
961 TCAACACCAACATATGAAATGAGCGGAGCAAGTGTATATGTCGCCACACTCTTGCTG 1020
1021 GAAATCTGTGCAAAATTTCTACATCTGGAGCCAGGCAAGTGAATTTCTGGTGAACCTCGGG 1080
1021 GAAATCTGTGCAAAATTTCTACATCTGGAGCCAGGCAAGTGAATTTCTGGTGAACCTCGGG 1080
1081 CATTTGACTGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCAACTCTCTGTAA 1140
1081 CATTTGACTGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCAACTCTCTGTAA 1140
1141 GGTAACAGATCAGAGATATGACCGTCCCTACAGCAATGTGGAAGAGGTGAGACTGGC 1200
1141 GGTAACAGATCAGAGATATGACCGTCCCTACAGCAATGTGGAAGAGGTGAGACTGGC 1200
1201 TTTCAATCCAGAGAGCTGAAATGCTCTCTGAGGTGACCACTCATCTACCATTA 1260
1201 TTTCAATCCAGAGAGCTGAAATGCTCTCTGAGGTGACCACTCATCTACCATTA 1260
1261 AGAATATTTCTGAAATGGGCTCAGTGGATTTCTCTGGGGTTTGGATGCTCTCACCGTA 1320
1261 AGAATATTTCTGAAATGGGCTCAGTGGATTTCTCTGGGGTTTGGATGCTCTCACCGTA 1320
1321 TGTCAATGAATCATCATCTGATGAGCAGGAGGAGCAACCTTTCCAGGGAGGT 1380
1321 TGTCAATGAATCATCATCTGATGAGCAGGAGGAGCAACCTTTCCAGGGAGGT 1380
1381 GTGAGGCGGTATTTGTGAAGCATCTGACACTGACGATCTTAGGCAACCTCTCTGAGGATG 1440
1381 GTGAGGCGGTATTTGTGAAGCATCTGACACTGACGATCTTAGGCAACCTCTCTGAGGATG 1440
1441 GGGTAGACCCATGAAAGGCAAGATTAGGAGAGCAGACCTAGTATATATTTTTCAGA 1500
1441 GGGTAGACCCATGAAAGGCAAGATTAGGAGAGCAGACCTAGTATATATTTTTCAGA 1500
1501 TTCCCTGACCTTGGCACTTAATCCGACACTTACATTTTCTGTAATATAA 1560
1501 TTCCCTGACCTTGGCACTTAATCCGACACTTACATTTTCTGTAATATAA 1560
1561 GTACTTATTAGGTAATAGAGGTTTGTATGCTATATATATATATCTTACCATCTTGAAGGGT 1620
1561 GTACTTATTAGGTAATAGAGGTTTGTATGCTATATATATATATCTTACCATCTTGAAGGGT 1620
1621 AGGTTTACCTGATAGCCAGAAATATCTAGACATTTCTATATCTTACAGGTAATATCTC 1680
1621 AGGTTTACCTGATAGCCAGAAATATCTAGACATTTCTATATCTTACAGGTAATATCTC 1680
1681 TTTAAACACCTATTTGTTTTCTATAAGCCATATTTTGGAGCACTAAAGTAAATGGC 1740
1681 TTTAAACACCTATTTGTTTTCTATAAGCCATATTTTGGAGCACTAAAGTAAATGGC 1740
1741 AAATTTGGACAGATATGAGGCTGAGGATTTGTTGATTTGTTGATTTGACAAATAA 1800
1741 AAATTTGGACAGATATGAGGCTGAGGATTTGTTGATTTGTTGATTTGACAAATAA 1800
1801 GCTAGACATTTTACCTTTGTTGCCACAGACATAACACTACCTCAGGAAAGCTGAGCTGC 1860
1801 GCTAGACATTTTACCTTTGTTGCCACAGACATAACACTACCTCAGGAAAGCTGAGCTGC 1860
1861 TTTAAGGCAACACAAACAAATCAGTGTACAGTATGGATGAATCTATGTTAAGCAT 1920
1861 TTTAAGGCAACACAAACAAATCAGTGTACAGTATGGATGAATCTATGTTAAGCAT 1920
1921 CTCAGATAAGGCCCAAGTTTTTATAGTTGCATCTCAGGGAGAAATTTTATAGGATGTTT 1980

```

Db 1921 CTCAGAAATAGCCAGTTTATAGTTGCACTCAGGGAAGAAAATTTATAGATGTTT 1980
 Qy 1981 ATGAGTTCTCCAAATAGATGTTTCTGCAATACATATAAAAAAAGGCGG 2040
 Db 1981 ATGAGTTCTCCAAATAGATGTTTCTGCAATACATATAAAAAAAGGCGG 2040
 Qy 2041 CCGC 2044
 Db 2041 CCGC 2044

RESULT 2

US-10-042-431-45
 ; Sequence 45, Application US/10042431
 ; Publication No. US20020182675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MCCARTHY, Sean A
 ; APPLICANT: BARNES, Thomas M
 ; APPLICANT: FRASER, Christopher C
 ; APPLICANT: SHARP, John D
 ; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
 ; PREVENTIVE, THERAPEUTIC, AND OTHER USES
 ; FILE REFERENCE: 10147-6U2
 ; CURRENT APPLICATION NUMBER: US/10/042,431
 ; CURRENT FILING DATE: 2001-10-25
 ; PRIOR APPLICATION NUMBER: US 09/333,159
 ; PRIOR FILING DATE: 1999-06-14
 ; PRIOR APPLICATION NUMBER: US 09/578,063
 ; PRIOR FILING DATE: 2000-05-24
 ; NUMBER OF SEQ ID NOS: 79
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 45
 ; LENGTH: 2044
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-042-431-45

Query Match 100.08; Score 2044; DB 14; Length 2044;
 Best Local Similarity 100.08; Pred. No. 0;
 Matches 2044; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GTCGACCCACGCGTCCGGGAAATGACAGAGAAATATGTAAGAGATTTTAAACCCAC 60
 Db 1 GTCGACCCACGCGTCCGGGAAATGACAGAGAAATATGTAAGAGATTTTAAACCCAC 60
 Qy 61 AAATCTTCTTACTTTAGATTTAGTTGTTATGTTGCGAGGAAATATGTAAGAGATTT 120
 Db 61 AAATCTTCTTACTTTAGATTTAGTTGTTATGTTGCGAGGAAATATGTAAGAGATTT 120
 Qy 121 GGACCATGTTGAAACCTTGTCAAGACAGTGGATGTTCTCACAGAAATGGAATGTGGC 180
 Db 121 GGACCATGTTGAAACCTTGTCAAGACAGTGGATGTTCTCACAGAAATGGAATGTGGC 180
 Qy 181 TTCTGATTTCTGTTGGCGTATATGTTCCAGAGAAATGTAATTCAGTACATATGCCAACTA 240
 Db 181 TTCTGATTTCTGTTGGCGTATATGTTCCAGAGAAATGTAATTCAGTACATATGCCAACTA 240
 Qy 241 AAGCTGTGCAACCCAGCAATTCATGATATTTAGTGAATCATCCAAATCAAGGCTATC 300
 Db 241 AAGCTGTGCAACCCAGCAATTCATGATATTTAGTGAATCATCCAAATCAAGGCTATC 300
 Qy 301 CTTGTGAGGAATATGAAGTCCCACTGAAGATGGGTATATCTTTCTGTTTAAACAGGATTC 360
 Db 301 CTTGTGAGGAATATGAAGTCCCACTGAAGATGGGTATATCTTTCTGTTTAAACAGGATTC 360
 Qy 361 CTCGAGGCTAGTGCACCTAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
 Db 361 CTCGAGGCTAGTGCACCTAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
 Qy 421 GCCTAGTTGGAGGTGCTAGCAACTGGATTTCCAACTCCGCAACAAATAGCCTGGGCTTCA 480
 Db 421 GCCTAGTTGGAGGTGCTAGCAACTGGATTTCCAACTCCGCAACAAATAGCCTGGGCTTCA 480

Qy 481 TTCTGCGAGATGCTGGTTTGTGATGGGAAACAGACAGGGGAAACGCTGCTC 540
 Db 481 TTCTGCGAGATGCTGGTTTGTGATGGGAAACAGACAGGGGAAACGCTGCTC 540
 Qy 541 GAAAAACAAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTTCACTATGATGAGA 600
 Db 541 GAAAAACAAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTTCACTATGATGAGA 600
 Qy 601 TGGCTAGGTTTGAACCTTCTGCGATGATAAATTTTATTTGCAAGAAAGGCGCCAGGAAA 660
 Db 601 TGGCTAGGTTTGAACCTTCTGCGATGATAAATTTTATTTGCAAGAAAGGCGCCAGGAAA 660
 Qy 661 AGATCTATTATGTCGCTGCTTCAAGGCAACCACTGGGCTTTATTTGCAATTTTCCACCA 720
 Db 661 AGATCTATTATGTCGCTGCTTCAAGGCAACCACTGGGCTTTATTTGCAATTTTCCACCA 720
 Qy 721 TGGCAGAGCTGGCTCAGAAAATCAAAATGTAATTTGCTTTAGCACCACTAGCACTGTTA 780
 Db 721 TGGCAGAGCTGGCTCAGAAAATCAAAATGTAATTTGCTTTAGCACCACTAGCACTGTTA 780
 Qy 781 AGCATGCAAAAAAGCCCGGACCAAAATTTGCTGTCAGATATGATGATCAAGGAT 840
 Db 781 AGCATGCAAAAAAGCCCGGACCAAAATTTGCTGTCAGATATGATGATCAAGGAT 840
 Qy 841 TGTGTTGCAAAAAAGAAATTTCTGATCAGACAGAAATTTCTCAGACAATTTGTTATTACC 900
 Db 841 TGTGTTGCAAAAAAGAAATTTCTGATCAGACAGAAATTTCTCAGACAATTTGTTATTACC 900
 Qy 901 TTTGTTGCGCAGGTGATTTCTGATCAGATTTGATGATGATGATGATGATGATGATGAT 960
 Db 901 TTTGTTGCGCAGGTGATTTCTGATCAGATTTGATGATGATGATGATGATGATGATGAT 960
 Qy 961 TCAACACCAAAATATGAACATGAGCCGAGCAAGTGTATATGCTGCCCACTCTTGTG 1020
 Db 961 TCAACACCAAAATATGAACATGAGCCGAGCAAGTGTATATGCTGCCCACTCTTGTG 1020
 Qy 1021 GAACTCTGTCAGAAATATTTCTACTGAGCCGAGCAGTGAATTTCTGGTGAATCTCCGG 1080
 Db 1021 GAACTCTGTCAGAAATATTTCTACTGAGCCGAGCAGTGAATTTCTGGTGAATCTCCGG 1080
 Qy 1081 CAATTTGACTGGGGGAGTGAGACCAAAAAATCTGGAATAATGCAATCAGCCAACTCTCTG 1140
 Db 1081 CAATTTGACTGGGGGAGTGAGACCAAAAAATCTGGAATAATGCAATCAGCCAACTCTCTG 1140
 Qy 1141 GGTACAGAGTCAAGATATGACGGTCCCTACAGCAATGTGAGCAGAGGTGAGGACTGGC 1200
 Db 1141 GGTACAGAGTCAAGATATGACGGTCCCTACAGCAATGTGAGCAGAGGTGAGGACTGGC 1200
 Qy 1201 TTTCAAAATCCAGAAAGCTGAAATGCTGCTCTCTGAGGTGACCAACCTCTCTACCAT 1260
 Db 1201 TTTCAAAATCCAGAAAGCTGAAATGCTGCTCTCTGAGGTGACCAACCTCTCTACCAT 1260
 Qy 1261 AGAATATCTGAAATGGGCTCAGTGGATTTCACTGGGGTTGGATGCTCTCAGCGTA 1320
 Db 1261 AGAATATCTGAAATGGGCTCAGTGGATTTCACTGGGGTTGGATGCTCTCAGCGTA 1320
 Qy 1321 TGTACAAATCAAAATCACTCCATCTGATGAGCAGAGAGAGCAACCTTTCCAGGAGCGT 1380
 Db 1321 TGTACAAATCAAAATCACTCCATCTGATGAGCAGAGAGAGCAACCTTTCCAGGAGCGT 1380
 Qy 1381 GTGAGCCGTATTTGGAAGCATCTGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
 Db 1381 GTGAGCCGTATTTGGAAGCATCTGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
 Qy 1441 GGGCTAGGACCCATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
 Db 1441 GGGCTAGGACCCATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
 Qy 1501 TTCCCTGCACTTGGCACTAAATCCGACACTTACATTTTACATTTTCTGTAATTTAA 1560
 Db 1501 TTCCCTGCACTTGGCACTAAATCCGACACTTACATTTTACATTTTCTGTAATTTAA 1560

QY 1561 GTACTTATTAGGTAAATAGAGGTTTGTATGCTATATATATATCTACCATCTTGAAGGT 1620
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1561 GTACTTATTAGGTAAATAGAGGTTTGTATGCTATATATATATCTACCATCTTGAAGGT 1620
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1621 AGGTTTACCTGATAGCCAGAAAATATCTAGACATCTCTATATATCTAGGTAAATCTC 1680
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1621 AGGTTTACCTGATAGCCAGAAAATATCTAGACATCTCTATATATCTAGGTAAATCTC 1680
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1681 TTTAAACACCTATCTCTTTCTTATAAGCCATATTTTGGAGCCTAAAGTAAATGCG 1740
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1681 TTTAAACACCTATCTCTTTCTTATAAGCCATATTTTGGAGCCTAAAGTAAATGCG 1740
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1741 AAATTGGGACAGATATTGAGGCTCGAGCTGTGGATTTTGTGATTTTGAACAAATAA 1800
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1741 AAATTGGGACAGATATTGAGGCTCGAGCTGTGGATTTTGTGATTTTGAACAAATAA 1800
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1801 GCTAGACATTTTACCTTTGTCACAGAGACATACACTACTCTAGGAGCTGAGCTGC 1860
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1801 GCTAGACATTTTACCTTTGTCACAGAGACATACACTACTCTAGGAGCTGAGCTGC 1860
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1861 TTTAAGGACAAACAAACAAATCAGTGTGTACAGTATGATGAAATCTATGTTAAGCAAT 1920
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1921 CTGAGATAGGCAAGTTTATAGTGTGATCTCAGGGAAGAAATTTTATAGGATGTT 1980
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1921 CTGAGATAGGCAAGTTTATAGTGTGATCTCAGGGAAGAAATTTTATAGGATGTT 1980
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1981 ATGAGTCTTCCATTAATGATCTCTGATTTACATTAATTAATTAATTAATTAATTA 2040
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1981 ATGAGTCTTCCATTAATGATCTCTGATTTACATTAATTAATTAATTAATTAATTA 2040
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 2041 CCGC 2044
Db |||||
QY 2041 CCGC 2044
Db |||||

RESULT 3

US-09-835-996A-20
; Sequence 20, Application US/09835996A
; Patent No. US20020142953A1
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Dennis
; APPLICANT: Loeb, Debra
; APPLICANT: Montgomery, Julie
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhao, Qing
; APPLICANT: Wehrman, Tom
; APPLICANT: Dumanac, Radoje
; APPLICANT: Ren, Feiyan
; APPLICANT: Qian, Xiahong
; APPLICANT: Wang, Dunrui
; TITLE OF INVENTION: MATERIALS AND METHODS RELATING TO LIPID METABOLISM
; FILE REFERENCE: 28110/35915A
; CURRENT APPLICATION NUMBER: US/09/835,996A
; CURRENT FILING DATE: 2001-04-16
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-11-17
; PRIOR FILING DATE: 2000-09-22
; PRIOR FILING DATE: 2000-09-22
; PRIOR FILING DATE: 2000-08-03
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20

; LENGTH: 2035
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (189)..(1415)
US-09-835-996A-20
Query Match 97.5%; Score 1993.8; DB 10; Length 2035;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1995; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 18 GGGAAATTCAGCAGGAAAATATCTGAAGAGTCTTTTAAACCCACAAATCTCTTACTTTA 77
Db GGGAAATTCAGCAGGAAAATATCTGAAGAGTCTTTTAAACCCACAAATCTCTTACTTTA 98
QY 78 GAATTAGTCTTACATTTGGCAGGAAAATAAATTAATGAGATTTGGACCATTTTGGAAACC 137
Db GAATTAGTCTTACATTTGGCAGGAAAATAAATTAATGAGATTTGGACCATTTTGGAAACC 158
QY 138 TTGTCAAGACAGTGGATTGTCTCACACAGAAATGGAATGTGGCTTCTGATTTCTGTGGCG 197
Db TTGTCAAGACAGTGGATTGTCTCACACAGAAATGGAATGTGGCTTCTGATTTCTGTGGCG 218
QY 198 TATATGTTCCAGAGAAATGTAATTCAGTACATATCCAACTAAAGCTGTGGACCCAGAA 257
Db TATATGTTCCAGAGAAATGTAATTCAGTACATATCCAACTAAAGCTGTGGACCCAGAA 278
QY 258 GCATTTCATGAATATTAGTGAATATCCTCAACATCAAGGCTATCCCTGTGAGGAATATGAA 317
Db GCATTTCATGAATATTAGTGAATATCCTCAACATCAAGGCTATCCCTGTGAGGAATATGAA 338
QY 318 GTCCAACTGAAGATGGGTATATCTCTTCTGTTTAAAGGATTCCTCGAGGCTTAGTGCAA 377
Db GTCCAACTGAAGATGGGTATATCTCTTCTGTTTAAAGGATTCCTCGAGGCTTAGTGCAA 398
QY 378 CCTAAGAAAGACAGGTTCCAGGCTGTGGTGTACTGAGCATGCGCTAGTTGGAGGTGCT 437
Db CCTAAGAAAGACAGGTTCCAGGCTGTGGTGTACTGAGCATGCGCTAGTTGGAGGTGCT 458
QY 438 AGCAACTGGATTTCCAACTGCGCAACATAGCTGGCTTCTATCTGCGAGATGCTGCT 497
Db AGCAACTGGATTTCCAACTGCGCAACATAGCTGGCTTCTATCTGCGAGATGCTGCT 518
QY 498 TTTGACCTGTGGATGGGAAACAGCAGGGGAAACGCTGTCTCGAAAACACAGACACTC 557
Db TTTGACCTGTGGATGGGAAACAGCAGGGGAAACGCTGTCTCGAAAACACAGACACTC 578
QY 558 TCCATAGACCAAGATGAGTTCTGGGCTTTCAGTTATGATGAGATGCTAGTTGACCTT 617
Db TCCATAGACCAAGATGAGTTCTGGGCTTTCAGTTATGATGAGATGCTAGTTGACCTT 638
QY 618 CCTGCACTGATAACTTTATTTTGCAGAAACGCGGCAAGAAAGATCTATTATGTGCGC 677
Db CCTGCACTGATAACTTTATTTTGCAGAAACGCGGCAAGAAAGATCTATTATGTGCGC 698
QY 678 TATTACAGGCGCACCACTAGGCTTTTATGCAATTTTCCACCATGCGAGAGTGGCTCAG 737
Db TATTACAGGCGCACCACTAGGCTTTTATGCAATTTTCCACCATGCGAGAGTGGCTCAG 758
QY 738 AAATCAAAATGATTTTGTCTTTAGCACCCATGCTTAAAGCATGCAAAAGCCCC 797
Db AAATCAAAATGATTTTGTCTTTAGCACCCATGCTTAAAGCATGCAAAAGCCCC 818
QY 798 GGGACCAAAATTTTGTGCTGCGCAGATATGATGATCAAGGATTTGTTGGCAAAAGAA 857
Db GGGACCAAAATTTTGTGCTGCGCAGATATGATGATCAAGGATTTGTTGGCAAAAGAA 878
QY 858 TTTCTGTATCAGACCAATTTCTCAGACAACTTTGTTATTTTACCTTTTGTGCGCAGGTT 917
Db TTTCTGTATCAGACCAATTTCTCAGACAACTTTGTTATTTTACCTTTTGTGCGCAGGTT 938
QY 918 CTTCATCAGATTTTGTAGTAATATCATGTTACTTCTGGGTGGAATTCACACCAACATATG 977

939	DB	CTTGTATCAGAAATTTGTAGTAATAATCATGTTACTTCTGGGTGGAATTCAAACCAACAATATG	998
978	QY	AACATGAGCCGAGCAAGTGTATATGTCTGCCACACATCTTCTGCTGGAACATCTGTGCAAAAT	1037
999	DB	AACATGAGCCGAGCAAGTGTATATGTCTGCCACACATCTTCTGCTGGAACATCTGTGCAAAAT	1058
1038	QY	ATTCTACACTGGAGCCAGGCAGTGAATTCCTGTGTGAACTCCGGGCATTTGACTCGGGGAGT	1097
1059	DB	ATTCTACACTGGAGCCAGGCAGTGAATTCCTGTGTGAACTCCGGGCATTTGACTCGGGGAGT	1118
1098	QY	GAGACCAAAATCTGGA AAAATGCAATCAGCCAACTCCTCTGTAAGGTACAGAGTCAGAGAT	1157
1119	DB	GAGACCAAAATCTGGA AAAATGCAATCAGCCAACTCCTCTGTAAGGTACAGAGTCAGAGAT	1178
1158	QY	ATGACGGTCCCTACAGCAATGTGAGCAGGAGGTGAGGACTGGCTTTCCAAATCCAGAAGAC	1217
1179	DB	ATGACGGTCCCTACAGCAATGTGAGCAGGAGGTGAGGACTGGCTTTCCAAATCCAGAAGAC	1238
1218	QY	GTGAAAATGTCCTCTCTGAGGTGACCAACCTCATCTACCATAGAANAATTCCTGAATGG	1277
1239	DB	GTGAAAATGTCCTCTCTGAGGTGACCAACCTCATCTACCATAGAANAATTCCTGAATGG	1298
1278	QY	GCTCACGTGGATTTTCATCTCGGGTTTGGATGCTCCTCACCGTATGTACAAATGAAATCATC	1337
1299	DB	GCTCATGTGGATTTTCATCTCGGGTTTGGATGCTCCTCACCGTATGTACAAATGAAATCATC	1358
1338	QY	CATCTGATGACGAGGAGGAGACCAACCTTTCCAGGGAAGGTGTGAGGCCGTATTGTGGA	1397
1359	DB	CATCTGATGACGAGGAGGAGACCAACCTTTCCAGGGAAGGTGTGAGGCCGTATTGTGGA	1418
1398	QY	AGCATCTGACACTGACCGATCTTAGGACCAACCTCTGAGGGATGGGGCTAGGACCCATGAA	1457
1419	DB	AGCATCTGACACTGACCGATCTTAGGACCAACCTCTGAGGGATGGGGCTAGGACCCATGAA	1478
1458	QY	GGCAGAAATTACGGAGCAGAGACCTAGTATACATTTTTCAGATTCCTCGCACCTGGGCAC	1517
1479	DB	GGCAGAAATTACGGAGCAGAGACCTAGTATACATTTTTCAGATTCCTCGCACCTGGGCAC	1538
1518	QY	TAAATCCGACACTTACATTTTACATTTTTCATTTTTCGTAAATTTAAAGTACTTATTAGGTAAAT	1577
1539	DB	TAAATCCGACACTTACATTTTACATTTTTCATTTTTCGTAAATTTAAAGTACTTATTAGGTAAAT	1598
1578	QY	AGAGGTTTTGTATGCTATTATATATTCTACATCTTGAAGGGTAGGTTTACCTGATAGC	1637
1599	DB	AGAGGTTTTGTATGCTATTATATATTCTACATCTTGAAGGGTAGGTTTACCTGATAGC	1658
1638	QY	CAGAAAATATCTAGACATTCCTATATCATTCAGGTAAATCTCTTTAAAAACACTATTGT	1697
1659	DB	CAGAAAATATCTAGACATTCCTATATCATTCAGGTAAATCTCTTTAAAAACACTATTGT	1718
1698	QY	TTTTTCTATAGCCATATTTTGGAGCATTAAAGTAAATGGCNAATTTGGACAGATATT	1757
1719	DB	TTTTTCTATAGCCATATTTTGGAGCATTAAAGTAAATTTGGCNAATTTGGACAGATATT	1778
1758	QY	GAGTCTCGGAGTCTGTGGATTAATGTTTGACTTTGACAAAATAAGCTAGACATTTTCACTT	1817
1779	DB	GAGTCTCGGAGTCTGTGGATTAATGTTTGACTTTGACAAAATAAGCTAGACATTTTCACTT	1838
1818	QY	TGTTGCCACAGAGNCAATAACACTACTCTCGGAGCTGAGCTGCTTTTAAGGACAAACAA	1877
1839	DB	TGTTGCCACAGAGNCAATAACACTACTCTCGGAGCTGAGCTGCTTTTAAGGACAAACAA	1898
1878	QY	CAAAATCAGTGTTCAGTATGGATGAAATCTATGTTAAGCATTCCTCAGAAATAAGGCCAAG	1937
1899	DB	CAAAATCAGTGTTCAGTATGGATGAAATCTATGTTAAGCATTCCTCAGAAATAAGGCCAAG	1958
1938	QY	TTTTATAGTTGCATCTCAGGGAGAAATTTTATAGGATCTTTATGAGTTCTCCATATAA	1997
1959	DB	TTTTATAGTTGCATCTCAGGGAGAAATTTTATAGGATCTTTATGAGTTCTCCATATAA	2018
1998	QY	TGCATTTCTGCATTACAT	2014

Db 2019 TGCATTCTGCATTACAT 2035

RESULT 4

US-09-835-996A-9

Sequence 9, Application US/09835996A

Patent No. US20020142953A1

GENERAL INFORMATION:

APPLICANT: Ballinger, Dennis

APPLICANT: Loeb, Debra

APPLICANT: Montgomery, Julie

APPLICANT: Tang, Y. Tom

APPLICANT: Zhou, Ping

APPLICANT: Goodrich, Ryle

APPLICANT: Liu, Chenghua

APPLICANT: Asundi, Vinod

APPLICANT: Zhao, Qing

APPLICANT: Wehrman, Tom

APPLICANT: Drmanac, Radoje

APPLICANT: Ren, Feiyan

APPLICANT: Qian, Xiahong

APPLICANT: Wang, Dunrui

TITLE OF INVENTION: MATERIALS AND METHODS RELATING TO LIPID METABOLISM

FILE REFERENCE: 28110/35915A

CURRENT APPLICATION NUMBER: US/09/835,996A

CURRENT FILING DATE: 2001-04-16

PRIOR APPLICATION NUMBER: US 60/197,137

PRIOR FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: US 09/714,936

PRIOR FILING DATE: 2000-11-17

PRIOR APPLICATION NUMBER: US 09/667,298

PRIOR FILING DATE: 2000-09-22

PRIOR APPLICATION NUMBER: US 09/631,451

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: US 09/598,042

PRIOR FILING DATE: 2000-06-20

NUMBER OF SEQ ID NOS: 45

SOFTWARE: PatentIn version 3.0

SEQ ID NO 9

LENGTH: 1384

TYPE: DNA

ORGANISM: Homo sapiens

US-09-835-996A-9

Query Match 65.7%; Score 1343.4; DB 10; Length 1384;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1344; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 18 GGGAAATTGCAGCAGGAAAAATATGTGAAGAGCTTTTAAACCCACAAATCTCTTCTACTTTA 77

Db 39 GGGAAATTGCAGCAGGAAAAATATGTGAAGAGCTTTTAAACCCACAAATCTCTTCTACTTTA 98

Qy 78 GAATTAGTCTTTACATTTGGCAGGAAAAATAAATGCAGATGTTGGACCATGTTGGAAACC 137

Db 99 GAATTAGTCTTTACATTTGGCAGGAAAAATAAATGCAGATGTTGGACCATGTTGGAAACC 158

Qy 138 TTGTCAAGCAGTGGATTGTCTCACACAGAAATGGAATGTGGCTTCTGATTCTCTGTGGCG 197

Db 159 TTGTCAAGCAGTGGATTGTCTCACACAGAAATGGAATGTGGCTTCTGATTCTCTGTGGCG 218

Qy 198 TATATGTTCCAGAGAAATGTGAATTCAGTACATATGCCAACTAAAGCTGTGGACCCAGAA 257

Db 219 TATATGTTCCAGAGAAATGTGAATTCAGTACATATGCCAACTAAAGCTGTGGACCCAGAA 278

Qy 258 GCATTCAATGAATATTAGTGAATATCATCCAACTCAAGGCTATCCCTGTGAGGAATATGAA 317

Db 279 GCATTCAATGAATATTAGTGAATATCATCCAACTCAAGGCTATCCCTGTGAGGAATATGAA 336

Qy 318 GTCGCAACTGAAGATGGGTATATCCCTTCTGTGTAAACAGGATTCCTCGAGGCCTAGTGCAA 377

Db 339 GTCGCAACTGAAGATGGGTATATCCCTTCTGTGTAAACAGGATTCCTCGAGGCCTAGTGCAA 398

Qy 378 CCTAAGAAGACAGGTTCCAGGCCTGTGGTGTACTGACAGCATGGCCTAGTTGGAGGTGCT 437

Db CCTAAGAGAGAGAGTTCAGGCTGTGGTGTACTGAGCATGCGCTAGTTGGAGGTCT 458
 QY AGCAACTGAGTTTCAACCTGCCCAACATAGAGCTGGCTTCACTTCGGAGATGCTGGT 497
 Db AGCAACTGAGTTTCAACCTGCCCAACATAGAGCTGGCTTCACTTCGGAGATGCTGGT 518
 QY TTTGACGTGTGGATGGGGAACAGCAGAGGGGAAACGCTTGTCTCAAAACACAGACACTC 557
 Db TTTGACGTGTGGATGGGGAACAGCAGAGGGGAAACGCTTGTCTCAAAACACAGACACTC 578
 QY TCCATAGACCAAGATGAGTTCTGGCTTTTCACTTATGATGAGATGGCTAGTTGACCTT 617
 Db TCCATAGACCAAGATGAGTTCTGGCTTTTCACTTATGATGAGATGGCTAGTTGACCTT 638
 QY CTGCAAGTATTAACCTTTATTTTGCAGAAACCGGCCAGGAAAGATCTATTATGTCGGC 677
 Db CTGCAAGTATTAACCTTTATTTTGCAGAAACCGGCCAGGAAAGATCTATTATGTCGGC 698
 QY TATTCACAGGGCAACCATGGGCTTTATGCAATTTTCCACCATGCGAGAGCTGGCTCAG 737
 Db TATTCACAGGGCAACCATGGGCTTTATGCAATTTTCCACCATGCGAGAGCTGGCTCAG 758
 QY AAAATCAAAATGATTTTGTCTTTAGCACCCATAGCACTGTTAAGCATGCAAAAGCCCC 797
 Db AAAATCAAAATGATTTTGTCTTTAGCACCCATAGCACTGTTAAGCATGCAAAAGCCCC 818
 QY GGGACCAAAATTTTGTCTGCCAGATATGATGATCAAGGATTTTGGCAAAAGAA 857
 Db GGGACCAAAATTTTGTCTGCCAGATATGATGATCAAGGATTTTGGCAAAAGAA 878
 QY TTTCTGTATCAGACAGATTTCTCAGACAACTTGTATTTTACCTTTTGGCCAGGTGATT 917
 Db TTTCTGTATCAGACAGATTTCTCAGACAACTTGTATTTTACCTTTTGGCCAGGTGATT 938
 QY CTGTATCAGATTTGATTAATATCATGTTTCTCTGGGTGATTTCAACCAACCAATATG 977
 Db CTGTATCAGATTTGATTAATATCATGTTTCTCTGGGTGATTTCAACCAACCAATATG 998
 QY AACATGAGCCGAGCAAGTGTATGCTGCCACACCTTCTGCTGGAACATCTGTGCAAAAT 1037
 Db AACATGAGCCGAGCAAGTGTATGCTGCCACACCTTCTGCTGGAACATCTGTGCAAAAT 1058
 QY ATTCTACACTGAGCCAGCAGTGAATTTCTGTTGAATCTCCGGGCACTTTGACTGGGGAGT 1097
 Db ATTCTACACTGAGCCAGCAGTGAATTTCTGTTGAATCTCCGGGCACTTTGACTGGGGAGT 1118
 QY GAGACCAAAATCTGGAAAAATGCAATCAGCCAACTCTGTAGGTACAGAGTCAGAGAT 1157
 Db GAGACCAAAATCTGGAAAAATGCAATCAGCCAACTCTGTAGGTACAGAGTCAGAGAT 1178
 QY ATGACGGTCCCTACAGCAATGTGGACAGAGGTGAGGAGGTGAGGAGTCCAGAGAC 1217
 Db ATGACGGTCCCTACAGCAATGTGGACAGAGGTGAGGAGGTGAGGAGTCCAGAGAC 1238
 QY GTGAAAAATCTGCTCTCTGAGGTGACCAACTCATCTACCAATAGATATTTCTGAAATGG 1277
 Db GTGAAAAATCTGCTCTCTGAGGTGACCAACTCATCTACCAATAGATATTTCTGAAATGG 1298
 QY GCTCAGCTGGAATTCATCTGGGGTTTGGATGCTCTCACCCTGATGTACAAATGAATCATC 1337
 Db GCTCAGCTGGAATTCATCTGGGGTTTGGATGCTCTCACCCTGATGTACAAATGAATCATC 1358
 QY CATCTGATGACAGAGAGAGACCA 1362
 Db CATCTGATGACAGAGAGAGACCA 1383

RESULT 5

US-09-759-1308-416
 ; Sequence 416, Application US/097591308
 ; Publication No. US2003002279A1
 ; GENERAL INFORMATION:

APPLICANT: Millennium Pharmaceuticals, Inc.
 APPLICANT: McCarthy, Sean A
 APPLICANT: Fraser, Christopher C
 APPLICANT: Sharp, John D
 APPLICANT: Barnes, Thomas S
 APPLICANT: Kirt, Susan J
 APPLICANT: Mackay, Charles R
 APPLICANT: Myers, Paul S
 APPLICANT: Leiby, Kevin R
 APPLICANT: Wrighton, Nicolas
 APPLICANT: Goodearl, Andrew
 APPLICANT: Holtzman, Douglas A
 TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
 TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
 TITLE OF INVENTION: USES
 FILE REFERENCE: MPI00-5350NMIM
 CURRENT APPLICATION NUMBER: US/09/759,130B
 PRIOR FILING DATE: 2002-09-16
 PRIOR APPLICATION NUMBER: US 09/479,249
 PRIOR FILING DATE: 2000-01-07
 PRIOR APPLICATION NUMBER: US 09/559,497
 PRIOR FILING DATE: 2000-04-27
 PRIOR APPLICATION NUMBER: US 09/578,063
 PRIOR FILING DATE: 2000-05-24
 PRIOR APPLICATION NUMBER: US 09/333,159
 PRIOR FILING DATE: 1999-06-14
 PRIOR APPLICATION NUMBER: US 09/596,194
 PRIOR FILING DATE: 2000-07-14
 PRIOR APPLICATION NUMBER: US 09/342,364
 PRIOR FILING DATE: 1999-06-29
 PRIOR APPLICATION NUMBER: US 09/608,452
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: US 09/393,996
 PRIOR FILING DATE: 1999-09-10
 PRIOR APPLICATION NUMBER: US 09/602,871
 PRIOR FILING DATE: 2000-06-23
 PRIOR APPLICATION NUMBER: US 09/420,707
 PRIOR FILING DATE: 1999-10-19
 NUMBER OF SEQ ID NOS: 460
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 416
 LENGTH: 1269
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-759-1308-416

Query Match 62.1%; Score 1269; DB 11; Length 1269;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 126 ATGTTGAAACCTTGTCAAGACAGTGGATTGTCTCACAGAAATGGAATGCGCTTCTG 185
 Db 1 ATGTTGAAACCTTGTCAAGACAGTGGATTGTCTCACAGAAATGGAATGCGCTTCTG 60

QY 186 ATCTGCTGGCTATATGTTCCAGAGAAATGGAATTCAGTACATATGCCAATTAAGCT 245
 Db 61 ATCTGCTGGCTATATGTTCCAGAGAAATGGAATTCAGTACATATGCCAATTAAGCT 120

QY 246 GTGACCCCAAGAGCAATTCATGAATATTAGTGAATATCATCAACATCAAGGCTATCCCTGT 305
 Db 121 GTGACCCCAAGAGCAATTCATGAATATTAGTGAATATCATCAACATCAAGGCTATCCCTGT 180

QY 306 GAGGAATATGAAGTCCAACTGAAGATGGGTATATCTTCTGTTTAAACAGGATTCCTGA 365
 Db 181 GAGGAATATGAAGTCCAACTGAAGATGGGTATATCTTCTGTTTAAACAGGATTCCTGA 240

QY 366 GGCCTAGTCAACCTTAAAGAGACAGGTTCCAGGCTGTGGTGTACTGAGCATGGCCTA 425
 Db 241 GGCCTAGTCAACCTTAAAGAGACAGGTTCCAGGCTGTGGTGTACTGAGCATGGCCTA 300

QY 426 GTTGGAGGTGCTAGCAACTGGATTTCCAACTGCCCAACAATAGCCTGGGCTTCAATCTG 485
 Db 301 GTTGGAGGTGCTAGCAACTGGATTTCCAACTGCCCAACAATAGCCTGGGCTTCAATCTG 360


```

Db      721  GGCACAAAAGAAATTTCTGTATCAGACAGAAATTTCTCAGACAACTTGTATTTACCTTTGT 780
Qy      906  GGCAGGTGATTTCTGTATCAGATTTGTAGTAATATCATGTATTCTTCTGGGTGATTCATCAAC 965
Db      781  GGCAGGTGATTTCTGTATCAGATTTGTAGTAATATCATGTATTCTTCTGGGTGATTCATCAAC 840
Qy      966  ACCAACAAATATGAACATGAGCCGAGCAAGTGTATATGCTGCCACACACTCTTGCTGGAACA 1025
Db      841  ACCAACAAATATGAACATGAGCCGAGCAAGTGTATATGCTGCCACACACTCTTGCTGGAACA 900
Qy      1026  TCTGTGCAAAATATTTTACACTGAGCCAGGCAAGTGTATATGCTGGTGAATCTCGGGGATTT 1085
Db      901  TCTGTGCAAAATATTTTACACTGAGCCAGGCAAGTGTATATGCTGGTGAATCTCGGGGATTT 960
Qy      1086  GACTGGGGGAGTGAGACAAAATCTGAAAAATGCAATCAGCAAACTCTGTAAGGTAC 1145
Db      961  GACTGGGGGAGTGAGACAAAATCTGAAAAATGCAATCAGCAAACTCTGTAAGGTAC 1020
Qy      1146  AGAGTCAGAGATATGAGCGTCCCTACAGCAATGTGGACAGAGGTGAGACTGGCTTTCA 1205
Db      1021  AGAGTCAGAGATATGAGCGTCCCTACAGCAATGTGGACAGAGGTGAGACTGGCTTTCA 1080
Qy      1206  AATCCAGAGAGCGTGAATATGCTCTCTGAGGTGACCAACCTCATCTACCAATAGAAAT 1265
Db      1081  AATCCAGAGAGCGTGAATATGCTCTCTGAGGTGACCAACCTCATCTACCAATAGAAAT 1140
Qy      1266  ATTCTGAAATGGGCTCAGTGATTTTCACTGGGGTTTGGATGCTCTCACCCTATGATAC 1325
Db      1141  ATTCTGAAATGGGCTCAGTGATTTTCACTGGGGTTTGGATGCTCTCACCCTATGATAC 1200
Qy      1326  AATGAAATCATCTGATGAGCAGGAGGAGACCAACCTTTCCAGGAGCGGTGTAG 1385
Db      1201  AATGAAATCATCTGATGAGCAGGAGGAGACCAACCTTTCCAGGAGCGGTGTAG 1260
Qy      1386  GCCGTATTTG 1394
Db      1261  GCCGTATTTG 1269

```

```

RESULT 7
US-09-811-825-1
; Sequence 1, Application US/09811825
; Patent No. US20020144297A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC
; MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES
; THEREOF
; FILE REFERENCE: CLO01170
; CURRENT APPLICATION NUMBER: US/09/811.825
; CURRENT FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1206
; TYPE: DNA
; ORGANISM: Human
US-09-811-825-1

```

```

Query Match      55.1%; Score 1126; DB 10; Length 1206;
Best Local Similarity 100.0%; Pred. No. 3e-298;
Matches 1126; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      273  AGTGAATCATCCAACTCAAGGTATCCCTGTGAGGAATATGAATCGCAACTGAAGAT 332
Db      76  AGTGAATCATCCAACTCAAGGTATCCCTGTGAGGAATATGAATCGCAACTGAAGAT 135
Qy      333  GGGTATATCCCTTCTGTAAAGAGTATCCCTGAGGCTAGTGCACCTAAGAGACAGGT 392
Db      136  GGGTATATCCCTTCTGTAAAGAGTATCCCTGAGGCTAGTGCACCTAAGAGACAGGT 195
Qy      393  TCCAGGCTGTGGTGTACTGCAGCATGGCTAGTGGAGGTGCTAGCAACTGATTTCC 452

```

```

Db      196  TCCAGGCTGTGGTGTACTGTGAGCATGGCTAGTTGGAGGTGCTAGCAACTGGAATTTCC 255
Qy      453  AACCTGCCCAACAATAGACCTTGGCTTCAATTTGGCAGATGCTGGTTTGAAGTGTGATG 512
Db      256  AACCTGCCCAACAATAGACCTTGGCTTCAATTTGGCAGATGCTGGTTTGAAGTGTGATG 315
Qy      513  GGGAACACGACGAGGGAACGGCTTCTCGAAAAACACAAAGACACTCTCCATAGACCAAGAT 572
Db      316  GGGAACACGACGAGGGAACGGCTTCTCGAAAAACACAAAGACACTCTCCATAGACCAAGAT 375
Qy      573  GAGTTCTGGGCTTTTCACTATGATGATGGCTAGTTGGTGTGACCTTCTCGCAGTGAATAC 632
Db      376  GAGTTCTGGGCTTTTCACTATGATGATGGCTAGTTGGTGTGACCTTCTCGCAGTGAATAC 435
Qy      633  TTTATTTTGCAGAAAAACGGGCGCAGGAAAAAGATCTATTATGCTCGGCTATTTCACAGGGCACC 692
Db      436  TTTATTTTGCAGAAAAACGGGCGCAGGAAAAAGATCTATTATGCTCGGCTATTTCACAGGGCACC 495
Qy      693  ACCATGGGCTTTTATGATGATTTTCCACCATGCGCAGAGCTGGCTCAGAAAAATCAAAATGTAT 752
Db      496  ACCATGGGCTTTTATGATGATTTTCCACCATGCGCAGAGCTGGCTCAGAAAAATCAAAATGTAT 555
Qy      753  TTTGCTTTAGCACCATAGACCACTGTGTAAGCATGCAAAAAAGCCCGGACCAAAATTTTG 812
Db      556  TTTGCTTTAGCACCATAGACCACTGTGTAAGCATGCAAAAAAGCCCGGACCAAAATTTTG 615
Qy      813  TTGCTGCCAGATATGATGATCAAGGGATTTGTTGGCAAAAAAGAAATTTCTGTATCAGAC 872
Db      616  TTGCTGCCAGATATGATGATCAAGGGATTTGTTGGCAAAAAAGAAATTTCTGTATCAGAC 675
Qy      873  AGATTTCTCAGACAACTGTTATTTAGCTTTGGCCAGGCTGATTTCTGTATCAGATTTGT 932
Db      676  AGATTTCTCAGACAACTGTTATTTACCTTTGTGGCCAGGCTGATTTCTGTATCAGATTTGT 735
Qy      933  AGTAATATCATGTTACTTCTGGGTGATTTCAACCAACCAATATGAACATGAGCCGAGCA 992
Db      736  AGTAATATCATGTTACTTCTGGGTGATTTCAACCAACCAATATGAACATGAGCCGAGCA 795
Qy      993  AGTGATATGCTGCCACACTCTTGTGGAACATCTGTGCAAAATATTTCTACACTGGAGC 1052
Db      796  AGTGATATGCTGCCACACTCTTGTGGAACATCTGTGCAAAATATTTCTACACTGGAGC 855
Qy      1053  CAGGAGTGAATTTCTGGTGAATCTCGGGCAATTTGACTGGGGGAGTGAGACCAAAATCTG 1112
Db      856  CAGGAGTGAATTTCTGGTGAATCTCGGGCAATTTGACTGGGGGAGTGAGACCAAAATCTG 915
Qy      1113  GAAAAATGCAATCAGCCAACTCTCTGTAAGGTACAGAGTACAGAGATATGAGCGTCCCTACA 1172
Db      916  GAAAAATGCAATCAGCCAACTCTCTGTAAGGTACAGAGTACAGAGATATGAGCGTCCCTACA 975
Qy      1173  GCAATGTGGAAGAGGCTCAGGACTGGCTTTCAAAATCAGAGAGCGTGAATAATCTGCTC 1232
Db      976  GCAATGTGGAAGAGGCTCAGGACTGGCTTTCAAAATCAGAGAGCGTGAATAATCTGCTC 1035
Qy      1233  TCTGAGGTGACCAACCTCATCTACCAATAAGAAATTTCTGAATGGGCTCAGTGGATTTTC 1292
Db      1036  TCTGAGGTGACCAACCTCATCTACCAATAAGAAATTTCTGAATGGGCTCAGTGGATTTTC 1095
Qy      1293  ATCTGGGTTTGGATGCTCTCAGCGTATGTACAAATGAATATCATCTGATGATGAGGAG 1352
Db      1096  ATCTGGGTTTGGATGCTCTCAGCGTATGTACAAATGAATATCATCTGATGATGAGGAG 1155
Qy      1353  GAGGAGACCAACCTTTCCAGGAGCGGTGTGAGGCGGTATTGTGAA 1398
Db      1156  GAGGAGACCAACCTTTCCAGGAGCGGTGTGAGGCGGTATTGTGAA 1201

```

```

RESULT 8
US-09-811-825-3
; Sequence 3, Application US/09811825
; Patent No. US20020144297A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.

```


; TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES
 ; FILE REFERENCE: CL001170
 ; CURRENT APPLICATION NUMBER: US/09/811,825
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 18554
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-811-825-3

Query Match 43.6%; Score 890.4; DB 10; Length 18554;
 Best Local Similarity 99.9%; Pred. No. 9.4e-233;
 Matches 891; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1126	AGCCAACTCTGTAAGGTACAGAGTCAGAGATATGAGGTCCTCTACAGCAATGTGGACAG	1185
DB	16283	AGCCAACTCTGTAAGGTACAGAGTCAGAGATATGAGGTCCTCTACAGCAATGTGGACAG	16342
QY	1186	GAGGTACAGACTGGCTTTCAAATCCAGAGACGTGAATGCTCTCTCTGAGGTGACCA	1245
DB	16343	GAGGTACAGACTGGCTTTCAAATCCAGAGACGTGAATGCTCTCTCTGAGGTGACCA	16402
QY	1246	ACCTCATCTACCAATAGAAATATCTGAATGGCTCAGCTGGATTTTCATCTGGGTTGG	1305
DB	16403	ACCTCATCTACCAATAGAAATATCTGAATGGCTCAGCTGGATTTTCATCTGGGTTGG	16462
QY	1306	ATGCTCTCTACCGTATGTACAAATGAATCATCTGATGAGGAGGAGGACCAACC	1365
DB	16463	ATGCTCTCTACCGTATGTACAAATGAATCATCTGATGAGGAGGAGGACCAACC	16522
QY	1366	TTTCCAGGACCGGTGTAGGCGGTATGTGAAGCATCTGACATGACATCTTAGGACA	1425
DB	16523	TTTCCAGGACCGGTGTAGGCGGTATGTGAAGCATCTGACATGACATCTTAGGACA	16582
QY	1426	ACCTCTCTGAGGATGGGCTAGACCCATGAGCCAGATTAACGGAGGAGGAGCCTAG	1485
DB	16583	ACCTCTCTGAGGATGGGCTAGACCCATGAGCCAGATTAACGGAGGAGGAGCCTAG	16642
QY	1486	TATACATTTTTCAGATTTCCCTGACATTTGGCACTAAATCCGACACTTACATTTT	1545
DB	16643	TATACATTTTTCAGATTTCCCTGACATTTGGCACTAAATCCGACACTTACATTTT	16702
QY	1546	TTTCTGTAAATTAAGTACTTATAGGTAATAGAGGTTTGTATGCTATATATATCT	1605
DB	16703	TTTCTGTAAATTAAGTACTTATAGGTAATAGAGGTTTGTATGCTATATATATCT	16762
QY	1606	ACCATCTTGAAGGTAGGTTTACCTGATAGCCAGAAATATCTAGACATTTCTATATC	1665
DB	16763	ACCATCTTGAAGGTAGGTTTACCTGATAGCCAGAAATATCTAGACATTTCTATATC	16822
QY	1666	ATTCAGTAAATCTCTTTAAACACCTATTGTTTCTTATAAGCCATATTTTGGAGCA	1725
DB	16823	ATTCAGTAAATCTCTTTAAACACCTATTGTTTCTTATAAGCCATATTTTGGAGCA	16882
QY	1726	CTAAAGTAAATAGCAATTTGGGACAGATATTCAGGTCCTGAGTCTGTGATTTGTTG	1785
DB	16883	CTAAAGTAAATAGCAATTTGGGACAGATATTCAGGTCCTGAGTCTGTGATTTGTTG	16942
QY	1786	ACTTTGACAAAATAAGCTAGACATTTTACCTTTGTTGCCACAGACATATACTACCTC	1845
DB	16943	ACTTTGACAAAATAAGCTAGACATTTTACCTTTGTTGCCACAGACATATACTACCTC	17002
QY	1846	AGGAAGCTGAGTCTTTTAAAGGACAAACAAACAAATCAGTCTTACGATGTGATGAA	1905
DB	17003	AGGAAGCTGAGTCTTTTAAAGGACAAACAAACAAATCAGTCTTACGATGTGATGAA	17062
QY	1906	TCTATGTAAAGCATTTCTCAGAAATAGGCGCAAGTTTATAGTTGCTATCTCAGGGAAGAAA	1965
DB	17063	TCTATGTAAAGCATTTCTCAGAAATAGGCGCAAGTTTATAGTTGCTATCTCAGGGAAGAAA	17122

QY 1966 TTTTATAGGATGTTTATGAGTCTCCCAATAAATGCAATCTTGCATTACATATAA 2017
 DB 17123 TTTTATAGGATGTTTATGAGTCTCCCAATAAATGCAATCTTGCATTACATATAA 17174
 RESULT 9
 US-09-880-107-3878
 ; Sequence 3878, Application US/09880107
 ; Patent No. US20020142981A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Horne, Darci T.
 ; APPLICANT: Vockley, Joseph G.
 ; APPLICANT: Scherf, Uwe
 ; APPLICANT: Gene Logic, Inc.
 ; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
 ; CURRENT APPLICATION NUMBER: US/09/880,107
 ; CURRENT FILING DATE: 2001-06-14
 ; PRIOR APPLICATION NUMBER: US 60/211,379
 ; PRIOR FILING DATE: 2000-06-14
 ; PRIOR APPLICATION NUMBER: US 60/237,054
 ; PRIOR FILING DATE: 2000-10-02
 ; NUMBER OF SEQ ID NOS: 3950
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3878
 ; LENGTH: 2481
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: Genbank Accession No. US20020142981A1 Z31690
 US-09-880-107-3878

Query Match 25.8%; Score 527.8; DB 10; Length 2481;
 Best Local Similarity 66.6%; Pred. No. 1e-133;
 Matches 751; Conservative 0; Mismatches 376; Indels 0; Gaps 0;

QY	240	AAAGCTGTGGACCCAGAGGATTCATGAATATAGTGAATATCCCAATCCTCAACATCAGGCTAT	299
DB	112	ACAGCTGTGGATCTCTGAAACAAACATGATGAGTGAATATCTCTTACTGGGGATTC	171
QY	300	CCCTGTGAGGAATATGAAGTGGCAACTGAAGATGGGTATATCTCTTCTGTTAAACAGGATT	359
DB	172	CCTAGTGAGGAATACCTAGTGTGACAGAGATGGATATATCTGTGCTTAAACCGAATT	231
QY	360	CCTGAGCCCTAGTGCACCACTAAGAACACAGGTTCCAGGCTGTGGTGTACTGACGAT	419
DB	232	CCTCATGGGAGGAAGAACCACTTCTGCAAAAGGTCCTCAACCAAGTGTCTTCTGCAACAT	291
QY	420	GGCCTAGTGTGAGGTGTGCTAGCACTGGATTTTCCAACTGCTCCCAACATAGCTGGGCTTC	479
DB	292	GGCTTGTGTCAGATTTCTAGTAACTGGGTCAACACCTTGCCACACAGCGCTGGGCTTC	351
QY	480	ATTCTGGCAGATGCTGGTTTGTGATGGGAAACAGCAGGGAACAGCGCTGTCT	539
DB	352	ATTCTGTGTGATGCTGGTTTGTGATGGGAAACAGCAGGGAACAGCGCTGTCT	411
QY	540	CGAAACACACAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTTCACTATGATGAG	599
DB	412	CGAAACACACAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTTCACTATGATGAG	471
QY	600	ATGGTAGTGTGATGCTGCTTCCATAGTAAACCTTTATTTTGGAGAAAACGGGCGACGAA	659
DB	472	ATGGCAGAAATATGACCTACCAAGCTTCCATTAATCTTCAATTAATAAATGCGGCAAGAA	531
QY	660	AAGATCTATTATGTCGGCTATTTACAGGCGCACCATGGCTTTTATTTGATTTTCCACC	719
DB	532	CAAGTGTATTATGTCGGCTATTTACAGGCGCACCATGGCTTTTATTTGATTTTCCACC	591
QY	720	ATGCCAGAGCTGGCTCAGAAAATCAAAATGTATTTTCTTTTAGCACCCTAGCCATGTT	779
DB	592	ATCCCTGAGCTGGCTAAAAGGATTAATGTTTTTTCCTGGCTGCTGCTGGCTTCGCTC	651

```

Qy 780 AAGCATGCAAAAGCCCGGACCAAAATTTTGTGTCGACATATGATGATCAAGGCA 839
Db 652 GCCTTCGTACTAGCCCTATGGCAATTAGGACGATACAGATCATCTCATTAAGGAC 711
Qy 840 TTGTTGGCAAAAGAAATTTCTGTATCAGACAGATTTCTCAGACAACTTGTTATTTAC 899
Db 712 TTATTTGGACAAAGAAATTTCTCCCGACAGTGGCTTTTGAAGTGGGTGAGTCCAC 771
Qy 900 CTTTGTGGCCAGGTGATTTCTGTATGATGATGATGATGATGATGATGATGATGAT 959
Db 772 GTTTGCACTCATGTCTATGAGGAGCTCTGTGGAAATCTCTGTTTCTCTGTGTGA 831
Qy 960 TTCAACCAACAATATGACATGACGCGGACAGTGTATATGCTGCCACACTCTTGCT 1019
Db 832 TTTAATGAGAGAAATTTAAATATGCTAGAGTGGATGATATATCAACACATTTCTCTGCT 891
Qy 1020 GGAACATCTGTGCAAAATTTCTACACTGGAGCCAGGACGATGATTTCTGGTAACCTCCGG 1079
Db 892 GGAATCTGTGCAAAATGTTTACACTGGAGCCAGGCTGTAAATTTCCAAAGTTTCAA 951
Qy 1080 GCATTTGACTGGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCCAACTCTCTGTA 1139
Db 952 GCCTTTGACTGGGAGGAGTGGCCAGAAATTTATTTTCAATACCAAGAGTTATCTCTCC 1011
Qy 1140 AGGTACAGAGTCAGAGATATGACGCTCCCTACAGCAATGTGGACAGGAGTCAAGGACTGG 1199
Db 1012 ACATACATGTCAGGACATGCTTGTGCGACTGCACTCTGGAGCGGGGTCAAGCTGG 1071
Qy 1200 CTTTCAATCCAGAGAGTGAATGCTGCTCTCTGAGTGAGCAACCTCATCTACCAT 1259
Db 1072 CTTGCAGATGTCAGGAGTCAATATCTTACTGACTCAGATCACCACACTTGGTGTCCAT 1131
Qy 1260 AAGAAATTTCTGAATGGGTCACGTGATTTCTATCTGGGTTTGGATGCTCCTCACCGT 1319
Db 1132 GAGAGCATTCGGATGGAGCTCTTGACTTCAATTTGGGCGCTGGATGCCCTTGAGG 1191
Qy 1320 ATGTACAAATGAATCATCTCATCTGATGACAGAGGAGGACCAACCT 1366
Db 1192 CTTTATATATAAATTTATTAATCTAATGAGGAATATCAGTGAAGCT 1238

```

RESULT 10

```

US-10-198-846-13431
; Sequence 13431, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13431
; LENGTH: 2858
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 2718, 2719, 2720, 2721, 2722, 2723, 2846, 2847, 2848,
; LOCATION: 2849, 2850, 2851, 2852, 2853, 2854, 2855, 2856, 2857, 2858,
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-13431

```

Query Match 25.8%; Score 527.8; DB 15; Length 2858;
Best Local Similarity 66.6%; Fred. No. 1.1e-133;

```

Matches 751; Conservative 0; Mismatches 376; Indels 0; Gaps 0;
Qy 240 AAAGCTGGGACCCAGAGCAATTCATGAATATTAGTGAATATCCCAACATCAAGGCTAT 299
Db 247 ACAGCTGTGGATCTCTGAAAACAAACATGATGTGTGTGAATTAATCTCTTACTGGGATTC 306
Qy 300 CCCTGTGAGGAATATGAAGTGCACACTGAAGATGGGTATATCTCTTTCTGTAAACAGGATT 359
Db 307 CCTAGTGTAGGAATACCTAGTTGAGACAGAGATGGATATATCTGTGCTCTTAACCGGAAT 366
Qy 360 CCTGAGGACCTAGTGTCAACCTTAAGAAGACAGGTTCCAGGCTGTGGTGTACTCCAGCAT 419
Db 367 CCTCATGGGAGGAGAACCAATCTGCAAGAGTCCCAACCAAGTTGTCTTCTTCTGCAACAT 426
Qy 420 GGCCTAGTTGAGGTGTAGCAACTGTGATTTCCAACTGCCCAACAATAGAGCTGGGCTTC 479
Db 427 GGCCTGTGTCGAGATTTCTAGTAACTGTGGTCAACAACTTGCACACAGCAGCTGGGCTTC 486
Qy 480 ATTCCTGGCAGATGCTGGTGTGAGTGTGGGAAACAGCAGGGAACGCTGGTCT 539
Db 487 ATTCCTGTGATGCTGGTGTGAGTGTGGGAAACAGCAGGGAACGCTGGTCT 546
Qy 540 CGAAACCAACAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTCAAGTTATGATGAG 599
Db 547 CGAAACCAACAGACACTCTCAGATTTCTCAGATGAAATTTCTGGGCTTTCAAGTTATGATGAG 606
Qy 600 ATGGCTAGGTTTGACCTTCTGAGTGAATGAGTGAATGAGTGAATGAGTGAATGAGTGAAT 659
Db 607 ATGGCAAAATATGACCTTACCAGCTTCCATTAACCTTCAATCTGAATTAACATGAGGCAAGAA 666
Qy 660 AAGATCTATTATGTCGGCTATTTCACAGGGCACCACTGAGTGTGGGCTTTATTTGCAAAAACG 719
Db 667 CAAGTGTATTATGTCGGCTATTTCACAGGGCACCACTGAGTGTGGGCTTTATTTAGCATTTTTCACAG 726
Qy 720 ATGGCAGAGCTGGCTCAGAAATCAAAATGATTTGCTTTAGCACCACATGAGCACTGTT 779
Db 727 ATCCCTGAGTGGCTTAAAGGATTAAGATGTTTTTGGCCCTGGTCTGGTGGCTTCCGTC 786
Qy 780 AAGCATGCAAAAGCCCGGACCAAAATTTTGTGTCGACAGATATGATGATCAAGGGA 839
Db 787 GCCTTCTGTACTAGCCCTATGCGCAATTAGGACGATTAACAGATCATCTCATTAAGGAC 846
Qy 840 TTGTTTGGCAAAAGAAATTTCTGTATCAGACAGATTTCTCAGACAACTGTTATTATTC 899
Db 847 TTATTTGGAGCAAAAGAAATTTCTCCCGACAGTGGGTTTTTGAAGTGGCTGGGTACCCAC 906
Qy 900 CTTTGTGGCCAGGTGATTTCTGATCAGATTTGTAGTAAATATCATGTTTACTTCTGGTGA 959
Db 907 GTTTGCACTCATGTCTACTAGGAGCTCTGTGGAATCTCTGTTTCTTCTGTGTGA 966
Qy 960 TTCAACCAACAATATGACATGACGCGGACAGTGTATGCTGCCCAACATCTTCTGCT 1019
Db 967 TTTAATGAGAGAAATTTAAATATGCTAGAGTGGATGATATATACACACATCTCTCTGCT 1026
Qy 1020 GGAACATCTGTGCAAAATATTCTACACTGGAGCCAGGAGTGAATTTCTGTGAACTCCGG 1079
Db 1027 GGAATCTCTGTGCAAAACATGTTACACTGGAGCCAGGCTGTTAAATTTCCAAAGATTTCAA 1086
Qy 1080 GCATTTGACTGGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCAACCTCTGTA 1139
Db 1087 GCCTTTGACTGGGGAAGCAGTGGCCAGAAATTTATTTTCAATCAACACAGATTAATCTCTCC 1146
Qy 1140 AGGTACAGAGTCAGAGATATGACGCTCCCTACAGCAATGTGGACAGGAGGTCAGGACTGG 1199
Db 1147 ACATACATGTCAGGACATGCTTGTGCGGACTGTCAGTCTGGAGCGGGGTCAAGACTGG 1206
Qy 1200 CTTTCAATCCAGAGAGCACTGGAATGCTGCTCTCTGAGGTGACCAACTCATCTACCAT 1259
Db 1207 CTTGCAGATGCTCTACGACGTCATATCTTACTGACTCAGATCACCACATTTGGTGTTCAT 1266
Qy 1260 AAGAAATTTCTGAAATGGGCTCACGTGGAATTTTCTGCGGTTTGGATGCTTCTCACCGT 1319
Db 1267 GAGAGCATTCGGATGGAGCATCTTGACTTCATTTGGGCGCTGGATGCCCTTGGAGG 1326

```


Db 475 ACAGCTGTGGATCCTGAAACAAACATGAATGTGAGTGAATATCTTACTGGGATTC 534
 QY 300 CCCTGTGAGGAATATGAAGTCGCACTGAAGATGGGTATATCTTCTGTAAACAGATT 359
 Db 535 CCTAGTGAGGAATACCTAGTTGAGACAGAGATGGATATATCTGTGCTTAAACGAATT 594
 QY 360 CTTGAGGCTTAGTGCACCTAAGACAGAGTTCAGGCTGTGAGTGTGTTACTGACGAT 419
 Db 595 CCTCATGGGAGGAGAACCACTTCTGACAAAGGTCCTCAAAACAGTGTCTTCTCTGCAACAT 654
 QY 420 GGCCTAGTGTGAGTGTGACCACTGATTTCCAACTGCTGCCAACCAATAGCTGGGCTTC 479
 Db 655 GCTTGTGCTGAGATCTTAGTAACTGGTCAAAACCTTCCAAACAGAGCTGGGCTTC 714
 QY 480 ATTCGTGAGATGCTGCTTTTGTAGTGTGATGGGGAACAGAGGGGAAACGCTGCT 539
 Db 715 ATTCCTGCTGATGCTGCTTTTGTAGTGTGATGGGGAACAGAGAGGAAATACCTGCT 774
 QY 540 CGAAACACAGACACTCTCATAGACCAAGATGATGTTCTGGGCTTTCAGTTATGATGAG 599
 Db 775 CGAAACATAGACACTCTCAGTTCTCAGAGTGAATCTGGGCTTTCAGTTATGATGAG 834
 QY 600 ATGCTAGTGTGACCTTCTGCACTGATGAATCTTATTTTGCAGAAACGGGCCAGGAA 659
 Db 835 ATGCAAAATATGACTTACCTACCTGCTTCCATTAATCTTCAATCTGAATAAACTGCCAAGAA 894
 QY 660 AAGATCTATTATGCTGGCTATTACAGGGCACCACTGGCTTTATTTGCAATTTTCCACC 719
 Db 895 CAAGTGATTATGCTGGGTCACTTCAAGGCCACCATATAGTTTATAGCATTTTTCACAG 954
 QY 720 ATGCCAGAGTGGCTCAGAAATCAAAATGATTTTGTCTTTAGCACCCATGACCTGTT 779
 Db 955 ATCCCTGAGCTGGCTAAAGGATTAATGTTTGTGCTGGCTTCTGGCTTCGCTC 1014
 QY 780 AAGCATGCAAAAGCCCGGGAACAAATTTTGTGCTGCAGATGATGATCAAGGA 839
 Db 1015 GCCTTCTGTACTAGCCCTATGGCCAAATTAGGACGATTACCATCATCTCATTAAAGGAC 1074
 QY 840 TTGTTTGGCAAAAGAAATTTCTGTATCAGACAGAGATTTCTCAGACAACTTGTATTATTC 899
 Db 1075 TTAATGGAGACAAAGATTTCTTCCAGAGTGGCTTTTGAAGTGGCTGGTCCACC 1134
 QY 900 CTTTGTGGCCAGGTGATCTTGTATCAGATTTGTAGTAATATCATGTTACTTCTGGGTGA 959
 Db 1135 GTTTTGCATCATGTCTAGAGAGCTCTGTGGAATCTCTGTTTCTTCTCTGTGGA 1194
 QY 960 TTCAACACCAATATGATGATCAGCGAGAGTGTATATGCTGCCACACTCTTGTCT 1019
 Db 1195 TTTAATGAGAGAAATTTAATATGCTAGAGTGAATGATATACACACATTTCTCTGCT 1254
 QY 1020 GGAAACATCTGTGCAAAATATTTCTACATGGAGCCAGGAGTGAATCTGGTGAATCCCG 1079
 Db 1255 GGAACTTCTGTGCAAAACATGTTACATGGAGCCAGGCTGTAAATTTCCAAAGTTTCAA 1314
 QY 1080 GCATTTGACTGGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCCCACTCTGTA 1139
 Db 1315 GCCTTTGACTGGGAGAGCAGTGCACCAAGATTAATTTTCAATACACAGATTAATCTCC 1374
 QY 1140 AGGTACAGAGTACAGATATGACCGTCCCTACAGCAATGTGGAGAGGTGAGGACTGG 1199
 Db 1375 ACATCAATGTGAAGGACATGCTTGTGCGACTGCTGAGTGGAGCGGGGTACGACTGG 1434
 QY 1200 CTTTCAATCAGAGAGGAGTGAATGCTGCTCTGAGGTGAGGACCACTCATCTACCAT 1259
 Db 1435 CTTGAGATGCTACGAGCTCAATATCTTACTGACTCAGATCAACCACTTGGTGTCCAT 1494
 QY 1260 AAGAAATTTCTGAATGGGCTCAGTGGATTTTCACTGGGGGTTTGAATGCTCTCCACCGT 1319
 Db 1495 GAGAGCATTCGGAAATGGAGCACTTGTGACTTCAATTTGGGGCTGGATGCCCTTGGAG 1554
 QY 1320 ATGTACAAATGAATCATCTGATGACAGAGGAGGAGCAACT 1366
 Db 1555 CTTTATAATAAATTTAATCTAATGAGGAATATCAGTGAAGCT 1601

RESULT 13

US-10-108-260A-1262
 ; Sequence 1262, Application US/10108260A
 ; Publication No. US20040005560A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HELIX RESEARCH INSTITUTE
 ; TITLE OF INVENTION: NO. US20040005560A1 full length cDNA
 ; FILE REFERENCE: H1-A0106
 ; CURRENT APPLICATION NUMBER: US/10/108,260A
 ; CURRENT FILING DATE: 2002-03-27
 ; NUMBER OF SEQ ID NOS: 5458
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 1262
 ; LENGTH: 2506
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-108-260A-1262

Query Match 22.6%; Score 462.2; DB 12; Length 2506;
 Best Local Similarity 67.0%; Pred. No. 1.1e-115;
 Matches 656; Conservative 0; Mismatches 323; Indels 0; Gaps 0;
 QY 388 CAGGTTCCAGGCTGTGTGTACTGCGAGCATGGCTAGTTGGAGGTGCTAGCAACTGGA 447
 Db 280 CTGGTCCCAACCACTGTTCTTCTGCAACATGGCTGCTGGCAGATTTCTAGTAACCTGG 339
 QY 448 TTTCCAACTGCCCAACAAATAGCTGGGCTTCACTTGGCGAGATGCTGGTGTGACGTGT 507
 Db 340 TCACAAACCTTGGCCAAACAGCAGCTGGGCTTCACTTCTGTGATGCTGGTGTGACGTGT 399
 QY 508 GGATGGGGAACAGCAGAGGGGAAACGCTGCTCTGCAAAACCAAGACACTCTCCATAGACC 567
 Db 400 GGATGGGGAACAGCAGAGGAAATACCTGCTCTGCAAAACCAAGACACTCTCAGTTCTC 459
 QY 568 AAGATGAGTTCTGGGCTTTCAGTTATGATGAGATGGCTAGGTTGACCTTCTGCGAGTGA 627
 Db 460 AGGATGAATCTGGGCTTTCAGTTATGATGAGATGGCAAAATATGACCTACCACTTCCA 519
 QY 628 TAACTTTTATTTTCAGAAACGGCCAGGAAAGATCTATTATGCTGGCTATTTCACAGG 687
 Db 520 TTAATCTCATTTCTGAATAAACTGGCCAGAAACAGTGTATTATGCTGGTCACTTCTCAG 579
 QY 688 GCACCAACATGGGCTTTATTTGCAATTTTCCACCAATGCGCAGAGCTGGCTCAGAAATCAAAA 747
 Db 580 GCACCACTATAGTGTATATAGCAATTTTTCACAGATCCCTGAGCTGCTTAAAGGATTAATA 639
 QY 748 TGTATTTTCTTACGACCCATAGCCACTGTTAGCATGCAAAAGCCCGGACCAAT 807
 Db 640 TGTATTTTGGCTGGGCTTCTGGGCTTCTGGCTTCTGTACTAGCCCTATGCGCAAT 699
 QY 808 TTTTGTCTGCTGCCAGATATGATGATCAAGGATTTGTTGGCAAAAGAAATTTCTGTATC 867
 Db 700 TAGGACGATTACCAATCATCTCTAAGGACTTATTGGAGACAAAGAAATTTCTTCCC 759
 QY 868 AGACAGATTTCTCAGACACTTGTATTATTTACCTTTTGGCCAGGTGATTTCTGATCAGA 927
 Db 760 AGAGTGGCTTTTGAAGTGGCTGGGTACCCACGTTTGCACCTCATCTACTAGGAGGC 819
 QY 928 TTTGTAGTAATATCATGTTACTTCTGGGTGATTCACCAACCAATATGAACATGAGCC 987
 Db 820 TCTGTGGAATCTCTGTTTCTTCTGTGTGATTTAATGAGAGAAATTTAATATGCTA 879
 QY 988 GAGCAAGTGTATATGCTGCCCACTCTTGTGTGGAACATCTGTGCAAAATATTCTTACT 1047
 Db 880 GAGTGGATGTATATACACACATTTCTCTGCTGGAACCTTCTGTGCAAAACATGTGTACT 939
 QY 1048 GGAGCCAGCAGTGAATTTCTGCTGAATCTCCGGCAATTTGACTGGGGAGTGAGACCAAAA 1107
 Db 940 GGAGCCAGCAGTGTGAATTTCCAAAGTTTCAAGCTTTGACTGGGGAGCAGTGCACAGA 999
 QY 1108 ATCTGGAAATATGCAATCAGCCAACTCTGTGTAAGGTACAGAGTCAGAGATATGACGCTCC 1167

Db	1000	ATTATTTTCATTACAACACAGGTTATCTCCACGTACAATGTGAAGGACATGCTTGTGC	1059
QY	1168	CTACAGCAATGTGGACAGGAGGTGAGACTTGGGCTTTCAAATCCAGAGACGTGAAATATGC	1227
Db	1060	CGACTCGAGTCTGGAGGGGGGTACGACTGGCTTGCAGATGTCTACGAGTCAATATCT	1119
QY	1228	TGCTCTCTGAGGTGACCAACCTCATCTPACCATTAAGATAATTCTCGAATGGCTCACGTGG	1287
Db	1120	TACTGACTCAGATCAACCAACTTGGTGTTCCATGAGAGCATTCGGAATGGAGCATCTTG	1179
QY	1288	ATTTCATCTGGGGTTTGGATGTCCTCACGTTATGTACAATGAAATCATCCATCTGATGC	1347
Db	1180	ACTTCATTTGGGGCTGGATGGCCCTTTGAGGGCTTTTATAATAAAATTTAAATCTAATGA	1239
QY	1348	AGCAGGAGGAGCAACCT	1366
Db	1240	GGAATATCATGTGAAAGCT	1258

```

RESULT 14
US-10-420-564-3
; Sequence 3, Application US/10420564
; Publication No. US20040001819A1
; GENERAL INFORMATION:
; APPLICANT: Bolen, Paul L.
; APPLICANT: Cihak, Paul, L.
; APPLICANT: Scharpf Jr., Lewis G.
; TITLE OF INVENTION: Recombinant Kid Pregastric Esterase and Methods for
; TITLE OF INVENTION: Its
; TITLE OF INVENTION: Production and Use
; FILE REFERENCE: IPF-0009
; CURRENT APPLICATION NUMBER: US/10/420,564
; CURRENT FILING DATE: 2003-04-22
; PRIOR APPLICATION NUMBER: US/10/043,665B
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 09/166,489
; PRIOR FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1411
; TYPE: DNA
; ORGANISM: Kid (Goat)
US-10-420-564-3

```

Qy	524	GGGAAACGCTGTCTCGAAAAACACAAGACACCTCTCCATAGACCAAGATGAGTTCCTGGGC	583
Db	421	AGGAAACACCTGTGGGCCCGAGGAAACATTTATACTATTACCAGAGCTCCCTGAAATCTGGGC	480
Qy	584	TTTCAGTTATGATGAGATGGCTAGGTTTGACCTTCCTGCGATGATATAAACTTTATTTTGCA	643
Db	481	TTTCAGCTTTGATGAAATGGCTGAATATGACCTTCATCTACAATTGATTTCACTTTAA	540
Qy	644	GAAACGGGCCGAGGAAAGATCTATTATGTGTGGCTATTTCACAGGGCACACCAATGGCTT	703
Db	541	GAGAACGAGGACGAAGAAGCTACACTATGTTGGCCATTTCCAAAGGACCAACCACTTGGTTT	600
Qy	704	TATTTGCAATTTTCCACCATGCGACAGCTGGCTCAGAAAAATCAAAATGATTTTTCCTTAGC	763
Db	601	TGTGGCTTTTACCAATCCCACTGCGCTGAATAATCGAAGTCTTCCATGCAATTAGC	660
Qy	764	ACCCATAGCCACTGTGTTAAGCATGCAAAAGCCCGGGACCAAAATTTTGTGTGCGCAGA	823
Db	661	CCCACTGCGCACAGTGAAGCACACCCAGAGCCTGTTTAACAAACTTGCACTTATTCCTCA	720
Qy	824	TATGATGATCAAGGGATTTGTTGCGCAAAAAAGAAATTTCTGATCAGACCAAGATTTCTCAG	883
Db	721	CTTCTCTTCCAAAGTATATTTGCTAACAAATGTTCTTACCCACACAAATTTTGTGAACA	780
Qy	884	ACAACCTGTTATTACCTTTTGGCCAGGTGATCTTGTATCAGATTTGTAGTAAATATCAT	943
Db	781	ATTTCTTGTGTTTGAAGTGTGCTCTCGTGAGACACTGGATGTCCTTTGTAAGAAATGCTT	840
Qy	944	GTTACTTCTGGGTGGATTCAACACCAACCAATATGAACATGAGCCGAGCAAGTGTATATGC	1003
Db	841	GTTTGGCAATTACTGGAGCTGACAAATAAAACTTCAACATGATGCTGTAGATGTATGT	900
Qy	1004	TGCCACACCTTTGCTGGAAACATCTGTGCAAAATATTTCACTGGAGCAGCAGTGAA	1063
Db	901	AGCACATAATCCACAGCAGGAGCTTCTGTTCAAAACATCTCCACTGGAGACAGGCTATTA	960
Qy	1064	TTCTGTGTAACCTCGGGCATTTGACCTGGGGAGTGAGACCAAAATCTGAAAAATGCAA	1123
Db	961	GTCCTGGAAANTCCAGCTTTTCACTGGGAGCGCTCAGTTGAGACCTAATGCAATTATAA	1020
Qy	1124	TCAGCCAACTCCTGTAAAGTACAGAGTCAAGATATGACGGTCCCTACAGCAATGTGGAC	1183
Db	1021	TCAGCCACCACTCCCATCTTACAATTTAAACAGCCATGAAATGTCCTCAATTCAGTATGGAG	1080
Qy	1184	AGGAGTCAAGACTGGCTTTCAAATCCAGAGACGTGAAAATGCTGCTCTCTCAGGTGAC	1243
Db	1081	TGCTGGCAGAGCCTGTGGCTGACCTCAGGATGTGACCTTTGCTTTCAAACTCTC	1140
Qy	1244	CAACCTCATCTACATAAGAATAATCCTGAATGGGCTCAGTGGATTTTCATCTGGGGTTT	1303
Db	1141	TAATCTCAATTCACCAACAGGAAATTCAAAATTAACAATCATCTGGAATTTATCTGGGCAAT	1200
Qy	1304	GGATGCTCTCACCGTATGTACAAATGAATCATCCATCTGATCAGCAGGAGAGA	1359
Db	1201	GGATGCACCTCAAGAAGTTTACAATGAAATATTTCTTTGATGGCAAAAGACAAA	1256

RESULT 15
US-10-312-088-7
; Sequence 7, Application US/10312088
; Publication No. US20030219862A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Cogswell, John P.
; APPLICANT: Kabnic, Karen S.
; APPLICANT: Lai, Ying-Ta
; APPLICANT: Martensen, Shelby A.
; APPLICANT: Murdock, Paul R.
; APPLICANT: Smith, Randall F.
; APPLICANT: Strum, Jay C.
; APPLICANT: Xiang, Zhaoying
; APPLICANT: Xie, Qing

```

RESULT 15
US-10-312-088-7
; Sequence 7, Application US/10312088
; Publication No. US20030219862A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Cogswell, John P.
; APPLICANT: Kabinic, Karen S.
; APPLICANT: Lai, Ying-Ta
; APPLICANT: Martensen, Shelby A.
; APPLICANT: Murdoch, Paul R.
; APPLICANT: Smith, Randall F.
; APPLICANT: Strum, Jay C.
; APPLICANT: Xiang, Zhongyong
; APPLICANT: Xie, Qing

```

```

; APPLICANT: Rizni, Safia K.
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50029
; CURRENT APPLICATION NUMBER: US/10/312,088
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: PCT/US01/19929
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,161
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/213,156
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1194
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-312-088-7

Query Match      22.2%; Score 454.2; DB 13; Length 1194;
Best Local Similarity 62.4%; Pred. No. 1e-113;
Matches 711; Conservative 0; Mismatches 428; Indels 0; Gaps 0;

QY 211 GAAATGTAATTCAGTACATATGCCAACTAAAGCTGTGGACCCAGAGCAATTCATGAATA 270
DB 44 GATCTATGATGTTATGACAGAAAGAAACATGCAACCCCTGAGCTAATATGAATA 103

QY 271 TTAGTGAATATCCAAATCAAGGCTATCCCTGTGAGGAATATGAAGTCGCACTGAAG 330
DB 104 TTAGCCAGATTAATTTCTTACTGGGGTTATCCTTATGAGAGATGATGTTTACACAAAG 163

QY 331 ATGGGTATATCCTTTCTGTTTAAAGGATTCCTCGAGGCTAGTGCAACCTAAGAGACAG 390
DB 164 ATGTTATATCCTTGGAAATTTATAGATTTCCATGGAAGAGATGCCCGAGGAGACAG 223

QY 391 GTTCCAGGCTGTGGTGTATCTGACAGATGGCTAGTGTGAGGTGCTAGCAACTGATTT 450
DB 224 CTCGAAAGCTGTGTGTTTTCAGCATGGCTTAATTCGATCTGCGAGTAACTGATTT 283

QY 451 CCACCTGCCCACATAGCTGGGCTTCAATCTGGCAGATGCTGGTGTGACGTGTGA 510
DB 284 GCAACCTGCCCACATAGCTGGGCTTCAATCTGGCAGATGCTGGTGTGACGTGTGA 343

QY 511 TGGGGAACAGCAGGGGAAACCGCTGTCTCGAAGAACAGACACTCTCCATAGACCAAG 570
DB 344 TGGGGAACAGCAGGGGAAACACTGTGTCCAGAAACACCTTAAATGTCACCGAAATCAC 403

QY 571 ATGAGTCTGGGCTTTTCAGTTATGATGAGATGGCTAGGTTTGACCTTCTGAGTGATA 630
DB 404 CGGATATCTGGGCTTTTCAGTTTGGATGAGATGGCTAAATGACCTTCCAGCCACATCA 463

QY 631 ACTTTATTTTCAGAAACCGGCGCAGGAAAGATCTTATTATGTCGGCTATTTCACAGGCA 690
DB 464 ATTTTATCATAGAGAAACCTGGACAGAGCGACTTACTACGTGGGCCACTCACAAGGCA 523

QY 691 CCACCTGGGCTTTATGCAATTTTCACCATGCGAGCTGGCTCAGAAATCAAATGT 750
DB 524 CCACCATAGCTTTTATAGCAATTTTCTACAAACCCAGACTGGCTAAAGATTAAGATAT 583

QY 751 ATTTTGTCTTTAGCACCCATAGCACTGTTAAGCATGCAAAAGCCCGGGACCAAAATTTT 810
DB 584 TTTTTCACCTGGCTCCAGTTGTCACAGTTAAATACACCCCAAGTCTATGAAAAAACTAA 643

QY 811 TGTGTGCGCAGATATGATGATCAGGGATTTGTTGGCAAAAGAAATTTCTGTATCAGA 870
DB 644 CAACCCCTTCCAGCGAGTAGTTAAGGTGTGTTGGTGACAAAATGTTCCACCCCTCATA 703

QY 871 CCAGATTTCTCAGACAACTTGTATTATTTACCTTTGTGGCCAGGCTGATTTCTTGATCAGATTT 930
DB 704 CATTTGTTTGACCAATTCATTTGCCACCAAGTGTGCAATCGAAGCTATTTCCGTCGTATTT 763

QY 931 GTAGTAATATCATGTTACTTCTGGGTGGATTTCAACACCAATATATGAACATGAGCCGAG 990

```

```

Db 764 GCAGCAACTTCCTATTCTCTGAGTGGATTTGATCCGCAAACTTTAAATATATGAGTCGCT 823
QY 991 CRAAGTGTATATGCTGCCACACATCTTGTCTGGGAACATCTGTGCAAAATATTTCTACACTGA 1050
Db 824 TGGATGTTTATTTGTGCACAAATCTCTGCGGGAACATCTGTTTCAAAATATGCTGCACTGG 883
QY 1051 GCCAGCAGTGAATTTCTGTGAACTCCGGGCAATTTGACTGGGGGAGTGAGACCAAAATC 1110
Db 884 CTCAGCTTTTACCACCTCTGATGAATTCAGAGCTTATGACTGGGGAATGACGCTGATAATA 943
QY 1111 TGGAAATATGCAATCAGCCCACTCTCTGTAAAGGTACAGAGTTCAGAGATATGACGGTCCCTA 1170
Db 944 TGAACATATACAATCAGAGTCTATCCCTTATATATGACCTGACTGCTGCAATGAAGTCCCTA 1003
QY 1171 CAGCAATGTGGACAGGAGGTCTAGGACTGGCTTTTCAAATCCAGAAGACGTGAAAATGCTGC 1230
Db 1004 CTGCTATTTGGGCTGTGGTGGACATGATGTCTCTGTAACACCCAGGATGTGGCCAGGATAC 1063
QY 1231 TCTCTGAGGTGACCAACCTCATCTACCATAGAATATTTCTGAAATGGGCTCAGGTGGATT 1290
Db 1064 TCCCTCAAATCAAGAGTCTTTTCAATCTTTAAGCTATTTGCCAGATTGGAACCACTTTGATT 1123
QY 1291 TCATCTGGGGTTTTGGATGCTCTCTCACCGTATGTACAAATGAATCAATCCATCTGAATCAG 1349
Db 1124 TTGCTGGGGCCCTCGATGCCCTCAACGGATGTACAGTGAATCATAGCTTTTAAATGAAG 1182

```

Search completed: February 19, 2004, 14:12:27
Job time : 747.122 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 04:41:57 ; Search time 4665.38 Seconds
(without alignments)
11127.553 Million cell updates/sec

Title: US-10-042-431-46
Perfect score: 1269
Sequence: 1 atgttgaaacctgttcaag.....gacggtgtgagccgtattg 1269

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: gb.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.ste.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg_hum.*
- 31: em.htg_inv.*
- 32: em.htg_other.*
- 33: em.htg_mus.*
- 34: em.htg_pln.*
- 35: em.htg_rtd.*
- 36: em.htg_man.*
- 37: em.htg_vrt.*
- 38: em_sy.*
- 39: em htgo_hum.*
- 40: em htgo_mus.*
- 41: em htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1122	88.4	1210	6	AX521798	AX521798 Sequence
2	1104.8	87.1	1272	6	AX477670	AX477670 Sequence
3	957	75.4	2577	10	BC031933	BC031933 Mus muscu
4	752	59.3	1116	6	AX521801	AX521801 Sequence
5	528.2	41.6	2382	9	AK091558	AK091558 Homo sapi
6	527.8	41.6	2481	6	AX411234	AX411234 Sequence
7	527.8	41.6	2481	9	HS1YACLY	Z31690 H.sapiens (
8	527.8	41.6	2481	9	HSU08464	U08464 Human lysos
9	527.8	41.6	2626	9	HS1AL	X76488 H.sapiens m
10	526.2	41.5	2493	6	BD094072	BD094072 Shear str
11	526.2	41.5	2493	9	HUMLPCHL	M74775 Human lysos
12	524.6	41.3	2586	9	BC012287	BC012287 Homo sapi
13	489.2	38.6	1378	6	A26689	A26689 Precursor o
14	489.2	38.6	1378	6	A26690	A26690 Precursor o
15	470	37.0	1365	9	HSGLR	X05997 Human mRNA
16	470	37.0	1367	6	A01046	A01046 H.sapiens m
17	470	37.0	1367	6	A12714	A12714 Pregastric
18	467.6	36.8	1336	6	A01157	A01157 R.norvegicu
19	467.6	36.8	1355	10	RNLIP	X02309 Rat mRNA fo
20	462.2	36.4	2506	9	AK096406	AK096406 Homo sapi
21	461	36.3	2358	10	MMLYACLY	Z31689 M.musculus
22	459.8	36.2	3144	10	S81497	L26319 Bos taurus
23	458	36.1	1417	4	BOVME	L26319 lysosomal a
24	454	35.8	1228	6	AX392821	AX392821 Sequence
25	448.4	35.3	1273	6	AX392824	AX392824 Sequence
26	446.8	35.2	1195	6	AX472752	AX472752 Sequence
27	444.2	35.0	1137	6	AR039022	AR039022 Sequence
28	444.2	35.0	1137	6	AR0392633	AR0392633 Sequence
29	444.2	35.0	1140	6	AR039023	AR039023 Sequence
30	444.2	35.0	1140	6	AR0392634	AR0392634 Sequence
31	444.2	35.0	1146	6	AR039024	AR039024 Sequence
32	444.2	35.0	1146	6	AR0392635	AR0392635 Sequence
33	444.2	35.0	1198	6	A57760	A57760 Sequence 5
34	444.2	35.0	1528	6	A57756	A57756 Sequence 1
35	444.2	35.0	1528	6	AR039021	AR039021 Sequence
36	444.2	35.0	1528	6	AR0392632	AR0392632 Sequence
37	444.2	35.0	1651	4	CFLIPASE	Y13899 Canis faml
38	441.8	34.8	1800	5	BC052131	BC052131 Danio rer
39	441	34.8	1137	6	A39301	A39301 Sequence 2
40	441	34.8	1140	6	A39303	A39303 Sequence 4
41	441	34.8	1146	6	A39305	A39305 Sequence 6
42	441	34.8	1528	6	A39300	A39300 Sequence 1
43	424.8	33.5	1260	6	AX645003	AX645003 Sequence
44	419.8	33.1	1086	6	AX710607	AX710607 Sequence
45	388.8	30.6	1048	6	A57758	A57758 Sequence 3

ALIGNMENTS

RESULT 1	AX521798	Sequence 1 from Patent WO0236731.	1210 bp	DNA	linear	PAT 05-OCT-2002
LOCUS	AX521798	Sequence 1 from Patent WO0236731.				
DEFINITION	AX521798					
ACCESSION	AX521798					
VERSION	AX521798.1	GI:23572854				
KEYWORDS		Homo sapiens (human)				
SOURCE		Homo sapiens				
ORGANISM		Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
REFERENCE		1				
AUTHORS		Xiao, Y.				
TITLE		Human lysosomal acid lipase				
JOURNAL		Patent: WO 0236731-A 1 10-MAY-2002;				

FEATURES		Bayer Aktiengesellschaft (DE)			
source		Location/Qualifiers			
		1..1210			
		/organism="Homo sapiens"			
		/mol_type="genomic DNA"			
		/db_xref="taxon:9606"			
CDS		1..1197			
		/note="unnamed protein product"			
		/codon_start=1			
		/protein_id="CAD52970.1"			
		/db_xref="GI:23572855"			
		/translation="MELKLLFRDVLFWAEMTRPLRKQSEIIHQSYFCEEEVEVATED GYLSVNRIPRGVLQPKTGSRPVLLQHLVGGASNWSLNNLSLFIADAGFDV WMNSGNASWRKHTLSIDQDFWAFSDENARFDPVAFINFLIKTQGEKIYYVG SQGTWGFAPSPWELAKIKMYPALAPIATVHAKSPGTRPFLLLPDMWIKGLPGKK EFLYQTFEFLRQVLYLGGVILDOICSNIMLLLGGPNTNNMNSRASVYAHTLAGTS VQNLWSQAVNSGEHLRADWGSSETXNLEKCNQPPVRYVRVDMVTPFAMWTGGQDL SNPDVKNLLSEVTNLIYHKNIPFWAFVDFIWLGLDAPHRWNEIHLKQOEBETNLSQG RCEAVL"			
BASE COUNT		338 a	264 c	296 g	312 t
ORIGIN					
Query Match		88.4%; Score 1122; DB 6; Length 1210;			
Best Local Similarity		100.0%; Pred. No. 0;			
Matches 112;		Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	148	AGTGAATCATCAACATCAAGGCTATCCCTGTGAGGAATATCAAGTCGCAACTGAAGAT	207		
Db	73	AGTGAATCATCAACATCAAGGCTATCCCTGTGAGGAATATCAAGTCGCAACTGAAGAT	132		
Qy	208	GGGTATATCTTTCTGTATACAGGATTCCTCGAGGCTAGTGCACCTAAGAGACAGCT	267		
Db	133	GGGTATATCTTTCTGTATACAGGATTCCTCGAGGCTAGTGCACCTAAGAGACAGCT	192		
Qy	268	TCCAGGCTGTGTGTACTGCAGCATGGCTAGTGTGGAGGTCTAGCAACTGGATTTCC	327		
Db	193	TCCAGGCTGTGTGTACTGCAGCATGGCTAGTGTGGAGGTCTAGCAACTGGATTTCC	252		
Qy	328	AACCTGCCCAACATAGCTGGGTTCAATCTGCGCAGATGCTGTTTACGCTGTGGATG	387		
Db	253	AACCTGCCCAACATAGCTGGGTTCAATCTGCGCAGATGCTGTTTACGCTGTGGATG	312		
Qy	388	GGGAACAGCAGGGGAAACGCTGTCTCGAAACACAAAGACATCTCCATAGACCAAGAT	447		
Db	313	GGGAACAGCAGGGGAAACGCTGTCTCGAAACACAAAGACATCTCCATAGACCAAGAT	372		
Qy	448	GAGTTCTGGGCTTTCAGTTATGATGAGATGGCTAGGTTTCACTTCTGTCAGTGATAAC	507		
Db	373	GAGTTCTGGGCTTTCAGTTATGATGAGATGGCTAGGTTTCACTTCTGTCAGTGATAAC	432		
Qy	508	TTTATTTTGCAGAAACGGGCGAGAAAGATCTATTATGCTGCTATTTCACAGGCGACC	567		
Db	433	TTTATTTTGCAGAAACGGGCGAGAAAGATCTATTATGCTGCTATTTCACAGGCGACC	492		
Qy	568	ACCATGGGCTTTATGCTATTTCACCATGCCAGAGCTGGCTCAGAAAATCAAAATGTAT	627		
Db	493	ACCATGGGCTTTATGCTATTTCACCATGCCAGAGCTGGCTCAGAAAATCAAAATGTAT	552		
Qy	628	TTTGCTTTAGCACCATTAGCCTGTTAAGCATGCACCAAGCCCGGAGCCCAATTTTG	687		
Db	553	TTTGCTTTAGCACCATTAGCCTGTTAAGCATGCACCAAGCCCGGAGCCCAATTTTG	612		
Qy	688	TTTGCTGCCAGATATGATGATCAAGGGATTTGTTGGCAAAAAGAAATTTCTGTATCAGACC	747		
Db	613	TTTGCTGCCAGATATGATGATCAAGGGATTTGTTGGCAAAAAGAAATTTCTGTATCAGACC	672		
Qy	748	AGATTTCTCAGACAACTGTTATTATCTTTGCGCCAGGTGATCTTTGATCAGATTTGT	807		
Db	673	AGATTTCTCAGACAACTGTTATTATCTTTGCGCCAGGTGATCTTTGATCAGATTTGT	732		
Qy	808	AGTAATATCATGTTACTTCTGGGTGGATTCAACACCAACATATGAACATGAGCCGACGA	867		
Db	733	AGTAATATCATGTTACTTCTGGGTGGATTCAACACCAACATATGAACATGAGCCGACGA	792		
LOCUS		AX477670			
DEFINITION		Sequence 8 from Patent WO0246418..			
ACCESSION		AX477670			
VERSION		AX477670.1			
KEYWORDS		GI:22216841			
SOURCE		Homo sapiens (human)			
ORGANISM		Homo sapiens			
REFERENCE		Griffin, J.A., Gandhi, A.R., Ramkumar, J., Tang, Y.T., Ding, L., Yue, H., Gietzen, K.J., Sapperstein, S.K., Honchell, C.D., Bruns, C.M., Duggan, B.M., Xu, Y. and Lee, S.			
TITLE		Lipid-associated molecules			
JOURNAL		Patent: WO 0246418-A 8 13-JUN-2002;			
FEATURES		Incyte Genomics, Inc. (US)			
source		Location/Qualifiers			
		1..1272			
		/organism="Homo sapiens"			
		/mol_type="genomic DNA"			
		/db_xref="taxon:9606"			
		/note="Incyte ID No: 7483978CB1"			
BASE COUNT		351 a	263 c	311 g	347 t
ORIGIN					
Query Match		87.1%; Score 1104.8; DB 6; Length 1272;			
Best Local Similarity		97.2%; Pred. No. 0;			
Matches 1143;		Conservative 0; Mismatches 12; Indels 21; Gaps 1;			
Qy	115	AAAGCTGGAGCCCGGAGGATTCATGAAATATTAGTAATATCCATCAACATCAAGGCTAT	174		
Db	94	AATGAAGTGAATTCCTGAGGTGTGGATGAATATCTAGTGAATATCCCAACATCAAGGCTAT	153		
Qy	175	CCCTGTGAGGATATGAAGTGCACACTGAAGATGGGTATATCCCTTTCTGTAAACAGGATT	234		
Db	154	CCCTGTGAGGATATGAAGTGCACACTGAAGATGGGTATATCCCTTTCTGTAAACAGGATT	213		
Qy	235	CCTCAGGCTTAGTGCAACCTTAAGAGACAGGTTCCAGGCTGTGGTGTACTCAGCAT	294		
Db	214	CCTCAGGCTTAGTGCAACCTTAAGAGACAGGTTCCAGGCTGTGGTGTACTCAGCAT	273		

Qy	61	ATTCGGTGGCTATATGTTTCAGAGAAATGGAATTCAGTACATATGCCAATTAAGCT	120
Db	172	ATTCGGTGGCTATATGTTTCAGAGAAATGGAATTCAGTACATATGCCAATTAAGCT	231
Qy	121	GTGGACCCAGAGAACATTCATGATATATAGTGAATATCATCCAAATCAAGGCTATCCCTGT	180
Db	232	GGGATCCAGAGAACATTCATGATATATAGTGAATATCATCCAAATCAAGGCTATCCCACT	291
Qy	181	GAGGAATATGAGTCCCACTGAAGTGGTATATCCCTTCGTTAAACAGGATTCCTCGA	240
Db	292	GAGGATATGAAGTTCGCAACAGAGATGGGTATCCCTTCGTTAAACAGGATTCCTCGG	351
Qy	241	GGCTAGTGCACCTAAGAGACAGGTTCCAGGCTGTGGTGTATCTAGCAGCATGCCCTA	300
Db	352	GGACACACACGGTTAAGAGAGAGGATCCAGGCCAGTGTGTACTGCAACATGCTCTT	411
Qy	301	GTGGAGGTGCTAGCAACTGGATTTCCAACTGCCCCCAACAAATAGCTGGGCTTCATTCG	360
Db	412	CTTGGGGATGCTAGCAACTGGATTTCCAACTGCCCCCAACAAATAGCTGGGCTTCATTCG	471
Qy	361	GCAAGATGCTGTTTTCAGAGTGGATGGGAAACAGCAGGGGAAACGCTGGTCTCGAAA	420
Db	472	GCAAGTGCAGGTTTGAATGTGGTGGGAAACAGCAGGGGAAACGCTGGTCTCGAAG	531
Qy	421	CACAAGACATCTCCATAGACCAAGATGAGTTCTGGGCTTTCAGTTATGATGAGTGGCT	480
Db	532	CACAAGACCTCTCCATAGATCAAGATGAGTTCTGGGCTTTCAGTTATGATGAGTGGCT	591
Qy	481	AGGTTGACCTCTCTGAGTGATTAATTTTTCAGAAACAGGGCCAGGAAAGATC	540
Db	592	AGGTTGACCTCTCTGAGTGATTAATTTTTCAGAAACAGGGCCAGGAAAGATC	651
Qy	541	TATTATGTCGGCTATTCACAGGGACCAACATGGGCTTTATGCAATTTCCACCATGCCA	600
Db	652	TATTATGTCGGCTATTCACAGGGACCAACATGGGCTTTATGCAATTTCCACCATGCCA	711
Qy	601	GAGCTGCTCAGAAATCAAAATGATTTTGTCTTATAGCACCACATAGCCTGTTAAGCAT	660
Db	712	GAGCTGCTCAGAAATCAAAATGATTTTGTCTTATAGCACCACATAGCCTGTTAAGCAT	771
Qy	661	GCAAAAGCCCGGACCAATTTTGTCTGCCAGATATGATGATCAAGGATGTTT	720
Db	772	GCAAGAGTCTGCTACCAAAATTTCTGCTGCCAGATATGATGATCAAGGATGTTT	831
Qy	721	GGCAAAAGAAATTTCTGTATCAGACCAAGTTTCTCAGACAACTTGTATTTACCTTCT	780
Db	832	GGCAGACAAGATTTTATACCACTAGATTTTTCAGGAGCTTTTATTTACCTTCT	891
Qy	781	GGCCAGGTGATTTCTTGATCAGATTTGTAGTAATATCATGTTACTTCTGGGTGGATTCAAC	840
Db	892	GGCCAGATGATTTCTGACCAAACTCTGCAGCAACATCATCTTCTCTGGGAGGTTTAAAC	951
Qy	841	ACCAACATATGAACATGAGCCGAGAGTGTATATCTGCCACACTCTGTGCTGGAACA	900
Db	952	ACCAACATATGAACATGAGCAGAGCAAAATGTGTATCTGCCACATACCTGTGGAACG	1011
Qy	901	TCTGTGCAAAATTTCTACACTGGAGCCAGGAGTGAATTTCTGTGAACTCCGGGCAATT	960
Db	1012	TCTGTGCAACATTTCTCACTGGAGCCAGGCTGATTTCTGGGAACTTCTGTGCTTT	1071
Qy	961	GACTGGGGAGTGAGACCAAAATTTCTGGAATAATGCAATGCGCACTCTGTAGGTAC	1020
Db	1072	GACTGGGGAGTGAGACCAAAATTTCTGGAATAATGCAATGCGCACTCTGTAGGTAC	1131
Qy	1021	AGAGTCAGAGATATGACGGTCCCTACAGCAATGTGGACAGGAGTCAAGACTGGCTTCA	1080
Db	1132	AAAGTTCAGATATGATGGTCCCAACAGCAATGTGGACTGAGGCTCAAGACTGGCTTCA	1191
Qy	1081	AATCCAGAGAGTGAATAATCTGCTCTCTGAGGTGACCAACCTCATCTACCATTAAGAT	1140
Db	1192	AATCCAGATGATGTGAATAATTTATTTCTGAAATGAACCACTCATCTACCATTAAGAT	1251
Qy	1141	ATTCCTGAATGGGCTCACTGGATTTTCACTCTGGGTTTGGATCTCTCACCGTATGTAC	1200
Db	1252	ATTCCTGAATGGGCTCACTGGATTTTCACTCTGGGTTTGGATCTCTCACCGTATGTAC	1311
Qy	1201	AATGAATATCATCATCTGATGACAGGAGGAGACCAACCTTTCCAGGACGGTGTGAG	1260
Db	1312	AATGAATATATACATCATCTGATGAAC---AGGAGCCCAACCTTTCCAGGAGACCTTCGAG	1369
Qy	1261	GCCTGATTG 1269	
Db	1369	GTCAATTTG 1377	
RESULT 4			
AX521801			
LOCUS			
DEFINITION			
Sequence 4 from Patent WO0236731.			
ACCESSION			
AX521801			
VERSION			
AX521801.1 GI:23572856			
KEYWORDS			
Homo sapiens (human)			
SOURCE			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE			
1			
Xiao, Y.			
Human lysosomal acid lipase			
Patent: WO 0236731-A 4 10-MAY-2002;			
Bayer Aktiengesellschaft (DE)			
FEATURES			
source			
1..1116			
/organism="Homo sapiens"			
/mol_type="genomic DNA"			
/db_xref="taxon:9606"			
BASE COUNT			
326 a 234 c 245 g 311 t			
ORIGIN			
Query Match			
59.3%; Score 752; DB 6; Length 1116;			
Best Local Similarity 97.4%; Pred. No. 2.5e-221;			
Matches 783; Conservative 0; Mismatches 0; Indels 21; Gaps 1;			
Qy	148	AGTGAATCATCCAAATCAAGGCTATCCCTGTGAGGAATATGAAGTCGCAACTGAAGAT	207
Db	73	AGTGAATCATCCAAATCAAGGCTATCCCTGTGAGGAATATGAAGTCGCAACTGAAGAT	132
Qy	208	GGGTATATCTCTGTTTAAACAGGATTCCTCGAGGCTAGTGCAACCTTAAGAACACAGGT	267
Db	133	GGGTATATCTCTGTTTAAACAGGATTCCTCGAGGCTAGTGCAACCTTAAGAACACAGGT	192
Qy	268	TCCAGGCTGTGGTGTACTGACAGATGGCTAGTGTGAGGTGTAGCAACTGGATTTC	327
Db	193	TCCAGGCTGTGGTGTACTGACAGATGGCTAGTGTGAGGTGTAGCAACTGGATTTC	252
Qy	328	RACCTGCCCAACAATAGCTGGCTTCATTTCTGCAGATGCTGGTTTTCAGCTGGATG	387
Db	253	RACCTGCCCAACAATAGCTGGCTTCATTTCTGCAGATGCTGGTTTTCAGCTGGATG	312
Qy	388	GGGAAACAGCAGGGGAAACGCTGTGTCGAAAAACAAAGACACTCTCCATAGACCAAGAT	447
Db	313	GGGAAACAGCAGGGGAAACGCTGTGTCGAAAAACAAAGACACTCTCCATAGACCAAGAT	372
Qy	448	GAGTTCCTGGCTTTCAGTTATGATGATGCTAGGTTTTCAGCTTCTCTGCAAGTATAAC	507
Db	373	GAGTTCCTGGCTTTCAGTTATGATGATGCTAGGTTTTCAGCTTCTCTGCAAGTATAAC	432
Qy	508	TTTATTTTGCAGAAAACGGGCGCAGGAAAGATCTATTATGTGCGCTATTTCACAGGCAAC	567
Db	433	TTTATTTTGCAGAAAACGGGCGCAGGAAAGATCTATTATGTGCGCTATTTCACAGGCAAC	492
Qy	568	ACCATGGGCTTTATTTGCAATTTTCCACCATGCGAGCTGGCTCAGAAATCAAAATGTAT	627
Db	493	ACCATGGGCTTTATTTGCAATTTTCCACCATGCGAGCTGGCTCAGAAATCAAAATGTAT	552

```

Qy 628 TTTCCTTAGCACCACATAGCCTGTTAAAGCATGCAAAAGAGCCCGGAGCAAAATTTTG 687
Db 553 TTTCCTTAGCACCACATAGCCTGTTAAAGCATGCAAAAGAGCCCGGAGCAAAATTTTG 612

Qy 688 TTGCTGCCAGATATGATGATCAAGGGATTTGTTGGCAAAAGAAATTTCTGTATCAGACC 747
Db 613 TTGCTGCCAGATATGATGATCAAGGGATTTGTTGGCAAAAGAAATTTCTGTATCAGACC 672

Qy 748 AGATTTCTCAGACAACTGTTTATTACCTTTGTCGGCCAGGTGATTTCTGTATCAGATTGT 807
Db 673 AGATTTCTCAGACAACTGTTTATTACCTTTGTCGGCCAGGTGATTTCTGTATCAGATTGT 732

Qy 808 AGTAATATCTGTTACTTCTGGGTGGATCAACACCAACCAATATGAACAT----- 857
Db 733 AGTAATATCTGTTACTTCTGGGTGGATCAACACCAACCAATATGAACATTAATCTCAT 792

Qy 858 -----GAGCCGAGCAGGTGATATGCTGCCACACTCTTGTGGAACATCTGTG 906
Db 793 GGTGTTGTACAGAGCCGAGCAGGTGATATGCTGCCACACTCTTGTGGAACATCTGTG 852

Qy 907 CAAATATTTCTACTGAGCCAG 930
Db 853 CAAATATTTCTACTGAGCCAG 876

RESULT 5
AK091558
LOCUS
DEFINITION
Homo sapiens cDNA FLJ34239 fis, clone FCBF3027755, highly similar
to LY5050MAL ACID LIPASE/CHOLESTERYL ESTER HYDROLASE PRECURSOR (EC
3.1.1.13).
ACCESSION
AK091558
VERSION
AK091558.1 GI:21749961
KEYWORDS
oligo capping; fis (full insert sequence).
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
REFERENCE
1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS
Ishibashi,T., Kanehori,K., Yosida,M., Watanabe,S., Ishida,S.,
Ono,Y., Hoteu,T., Hirooka,S., Murakawa,K., Takiguchi,S.,
Kusano,J., Watanabe,M., Fujimori,K., Tanai,H., Ishida,M.,
Yamashita,H., Chiba,Y., Sugiyama,T., Irie,R., Otsuki,T., Sato,H.,
Ota,T., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Matsuo,K.,
Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Watanabe,M.,
Takahaashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,
Sugano,S., Nagahara,K., Maehuo,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2982)
Isogai,T. and Yamamoto,J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, PLJ Project (HRI Team); 2-6-7
Kazusa-Kanetari, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
Location/Qualifiers
1 .2982
/organism="Homo sapiens"
/mol_type="cDNA"
/db_xref="taxon:9606"
/clone="FCBF3027755"
/tissue_type="brain"
/clone_lib="FCBF3"
/db_stages="fetal"
FEATURES
source

```

```

BASE COUNT      847 a      596 c      615 g      924 t
ORIGIN

Query Match      41.6%; Score 528.2; DB 9; Length 2982;
Best Local Similarity 65.4%; Pred. No. 7.8e-152;
Matches 792; Conservative 0; Mismatches 413; Indels 6; Gaps 1;

37  CACAGAAATGGAATGCGCTTCTGATTTCTGGTGGCGTATATGTTCCAGAGAAATGTGAAT 96
672 CACAGAAATGGAATGCGGTCTTGGGGTGGTGGTCTGTTGGTCTCTGACCCCTGCAT 731

97  TCAG-----TACATATGCCAACTAAAGCTGTGGACCCAGAGAACTTCAATGAATTAAGT 150
732 TCTGAGGGGTCTGGAGGGGAAATGACAGCTGTGGATCTCTGAAACAAACATGAATGTAGT 791

151 GAAATCATCCAAACATCAAGGGCTATCCCTGTGAGGAATATGAAGTCGCAACTGGAAGATGGG 210
792 GAAATTAATCTTACTGCGGGATTCCTAGTAGGGAATACCTAGTTGAGACAGAGATGGA 851

211 TATATCCTTTCTGTAAACAGGATTCCTCGAGCCCTAGTGCAACCTTAAGAACAGAGTTCC 270
852 TATATTTCTGTGCTTTAAACGGAATTCCTCATGGGAGAGAAACCAATTTCTGACAAAGGTCC 911

271 AGSCCTGTGGTGTACTGACAGCATGSCCTAGTTGGAGGTGTAGCAACTGTGATTTCCAAAC 330
912 AAACAGTTGCTTCTCTGCAACATGGCTTGTGGCAGATCTTAGTAATCTGGGTCAACAAC 971

331 CTGCCCAACAATPAGCTGGGCTTCATTTGGCAGATGCTGGTTTGGAGTGTGGATGGGG 390
972 CTTGCCAGCAGCAGCGCTGGGCTTCATTTCTGATGCTGGTTTGGAGTGTGGATGGGC 1031

391 AACACGAGGGGAAACGCGCTGTCTCGAAACACACAGACACTCTCCATAGACCAAGATCAG 450
1032 AACAGCAGAGGAAATACCTGGTCTCGGAAACATCAGACACTCTAGTTTCTCAGAGTGA 1091

451 TTCTGGGCTTTTCAGTTTATGATGAGATGGCTAGTTTGAACCTTCTCGCAGTGAATAACTTT 510
1092 TTCTGGGCTTTTCAGTTTATGATGAGATGGCAAAATATGACCTACCAGCTTCCATTAACCTC 1151

511 ATTTTCAGAAACCGGCGCAGGAAAGATCTATTTATGTCGGCTATTTCACAGGGCACCACC 570
1152 ATTTGAAATAAACCTGGCCAGAAACAGGTATTTATGTTGGGTTCATTTCTCAGAGGCCACT 1211

571 ATGGGCTTTATTCGCAATTTTCCACCATTGCGCAGAGCTGGCTCAGAAATCAAAATGTATTTT 630
1212 ATAGTTTTTATAGCATTTTCAGATCCCTGAGCTGGCTTAAAGGATTAAGATGTTTTTT 1271

631 GCTTTAGCACCCATAGCCACTGTTTAAGCATGCAAAAGCCCGGAGCAAAATTTTGTG 690
1272 GCGCTGGGTCTGTGGCTTCCGCTCTGTACTAGCCCTATGGCCAAATTAGGACGA 1331

691 CTGCCAGATATGATGATCAAGGGATTTGTTGGCAAAAGAAATTTCTGTATCAGACCCAGA 750
1332 TTACCAAGATCATCTCATTAAGACATTTTGGAGACAAAGAAATTTCTCCCAAGAGTGG 1391

751 TTCTCAGACAACTGTTTATTTACCTTTGGCCAGGTGATTTCTTGATCAGATTTGTAGT 810
1392 TTTTGAAGTGGCTGGGTACCCACCGTTTGCACTCATGTCACTCAAGAGAGCTCTGTGGA 1451

811 AATATCATGTTACTTCTGGTGGTGTTCACCAACAATATGAACATGAGCCGAGCAAGT 870
1452 AATCTCTGTTTTCTGTGTGGATTTTAAATGAGAAATTTAAATATGTTAGATGGAT 1511

871 GTATATGTGCCACACACTCTTCTCTGGAACATCTGTGCAAAATATTCTACCTGGAGCCAG 930
1512 GTATATACAAACATCTCTCTGCTGGAACCTTCTGTGCAAAACATGTTACCTACCTGGAGCCAG 1571

931 GCAGTGAATTTCTGGTGAATCTCGGGCATTTGCTGGGGAGTGAGACCAAAATCTGGAA 990
1572 GCTGTTAAATTCAAAAGTTTCAAGCCCTTTGACTGGGAGAGCAGTGCCAAAGATCATTTT 1631

991 AAATGCAATCAGCCAACTCTCTGTAAGGTACAGAGTACAGAGTATGACGGTCCCTACAGCA 1050

```

Db 1632 CATTACACAGAGTTATCTCCACATACATATGTGAAGGACATGCTTTGGCCGACATGCA 1691
 QY 1051 ATGTGGACAGAGGTTCAGGACTGGCTTTCAAAATCCAGAGACGTGAATAATGCTCTCT 1110
 Db 1692 GTCTGGAGCGGGGTTCAGGACTGGCTTTCAGATGCTACGAGCTCAATATCTTACTGACT 1751
 QY 1111 GAGGTGACCACTCATCTACATGAATATCTTGAATGGCTCAGTGGATTCATC 1170
 Db 1752 CAGATCACCACTTGGTGTTCATGAGAGCATTCGGGAATGGAGCATCTTGAATTCATT 1811
 QY 1171 TGGGGTGTGGATGCTCCTCACCGTATGTACATGAATCATCCATCTGATCAGCAGAG 1230
 Db 1812 TGGGGCTGGATGCTCCTGGAGGCTTTATAATAAAATTAATCTAATGAGGAATAT 1871
 QY 1231 GAGACCACT 1241
 Db 1872 CAGTGAAGCT 1882

RESULT 6
 AX411234 2481 bp DNA linear PAT 14-JUN-2002

LOCUS
 DEFINITION Sequence 3881 from Patent WO0229103.
 ACCESSION AX411234
 VERSION AX411234.1 GI:21443939
 KEYWORDS

SOURCE Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 1. Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J. G.
 Gene expression profiles in liver cancer
 Patent: WO 0229103-A 3881 11-APR-2002;
 GENE LOGIC INC (US)

FEATURES
 source
 1. 2481
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /notes="EMBL/GenBank Accession No. Z31690"
 BASE COUNT 692 a 477 c 513 g 799 t
 ORIGIN

Query Match 41.6%; Score 527.8; DB 6; Length 2481;
 Best Local Similarity 66.6%; Pred. No. 1e-151;
 Matches 751; Conservative 0; Mismatches 376; Indels 0; Gaps 0;

QY 115 AAAGCTGTGACCCAGAGCATTCATGAATATTAGTGAATCATCCAAATCAAGGCTAT 174
 Db 112 ACAGCTGTGATCTCTGAAACAAATGATGTGATGAATATCTTACTGGGATTC 171
 QY 175 CCCTGTGAGGAATATGAAGTCGCAACTGAAGATGGGTATATCCTTTCTGTTACAGGATT 234
 Db 172 CCTAGTGAGGAATACCTAGTTTGACAGAGAGATGATATATCTGTGCTTAAACCGAAT 231
 QY 235 CCTCGAGCCTAGTGCACTTAAGACAGAGTTCCAGGCTCTGTGTTTACTGACGAT 294
 Db 232 CCTATGGGAGGAGAACCAATCTGACAAAGGTCCTCCAAACAGTTCTTCTGCAAT 291
 QY 295 GGCCTAGTTGAGGTGTAGCAACTGGAATTCACAACTGCCCAACAAATAGCCTGGGCTTC 354
 Db 292 GGCCTGTGGCAGATTCAGTAACTGGGTACAAACCTTGCCCAACAGAGCCTGGGCTTC 351
 QY 355 ATTTGGCAGATGCTGTTTGGATGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 414
 Db 352 ATTTGCTGATGCTGTTTGGATGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 411
 QY 415 CGAAACACAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTCACTATGATGAG 474
 Db 412 CGAAACATAGACACTCTCAGTTCTCAGGATGAATCTTGGGCTTTCACTATGATGAG 471

QY 475 ATGGCTAGTGTGGCTTCTCTGAGTATAAATTTATTTTGCAGAAACGGCCAGAA 534
 Db 472 ATGGCAAAATATGACTACAGCTTCCATTAACCTTCTGAATAAACTGGCCAGAA 531
 QY 535 AAGATCTATATGTGGGTATTTCAAGGACCAACCTTGGGCTTTATTTGCATTTTCCACC 594
 Db 532 CAAGTGATATATGTGGGTCAATTTCAAGGACCAACCTTGGGCTTTATTTGCATTTTCAAG 591
 QY 595 ATGCCAGAGCTGGCTCAGAAAATCAAAATGTATTTTGTCTTTAGCAACCATGACCACTGTT 654
 Db 592 ATCCCTGAGCTGGCTAAAGGATTAATTTTGTCTTTTGGCTGGGCTGCTGGCTTCCGTC 651
 QY 655 AAGCATGCAAAAGCCCGGACCAAAATTTTGTGTGTCGCGAGATATGATGATCAAGGA 714
 Db 652 GCCTTCTGATAGCTTATGCCCATAATTTAGGACGATTTACCAAGATCATCTCAATAGGAC 711
 QY 715 TTGTTTGGCAAAAGAAATTTCTGTATCAGACGAGATTTCTCAGACAACTTTCTTATTATAC 774
 Db 712 TTATTTGGAGACAAAGAAATTTCTCCAGAGTGGCTTTTGAAGTGGCTGGGTACCCAC 771
 QY 775 CTTTGTGGCCAGGTGATTTCTGATCAGATTTGTAGTAATATCATGTTACTTCTGGTGA 834
 Db 772 GTTTGCACTCATGTCTACTGAAGAGCTCTGTGAAATCTCTGTTTCTTCTGTGTGA 831
 QY 835 TTCAACACCAACATATGACATGAGCGGAGCAAGTGTATATGCTGCCACACTCTTGTCT 894
 Db 832 TTTATGAGAGAAATTTTAAATATGCTAGAGTGAATATATATACACACATTTCTCTGCT 891
 QY 895 GGAACATCTCTGCAAAATATTTTACACTGGAGCGAGGAGTGAATTTCTGGTGAACCTCCGG 954
 Db 892 GGAATCTCTGCAAAACATGTTTACACTGGAGCGAGCTGTTAAATTTCCAAAGTTTCAA 951
 QY 955 GCATTTGACTGGGGGTGAGACCAAAATCTGAAAATGCAATCAGCCAACTCTCTGTA 1014
 Db 952 GCCTTTGACTGGGGAAGCAGTGCCTCAAGAAATTTTTCATTAACCAAGATTTCTCTCC 1011
 QY 1015 AGGTACAGAGTACAGATATGACGGTCCCTACAGCAATGTGGACAGAGGTGAGGCTGG 1074
 Db 1012 ACATACAAATGTAAGGACATGCTTTGTGCCGACTGCTGAGCGGGGTACAGACTGG 1071
 QY 1075 CTTTCAATCCAGAGACGTGAATATGCTCTCTGAGGTGAGCAACCTCATCTACCAT 1134
 Db 1072 CTTGAGATGTTACGAGCTCAATATCTTACTGACTAGATCACCACCTTGGTGTCCAT 1131
 QY 1135 AAGAATATTCCTGAATGGGCTCAGTGGATTTCTATGCGGGTTTGGATGCTCTCCCGT 1194
 Db 1132 GAGAGCTTCGGATCGGAGCATCTGACTTTCATTTGGGGCTGGATGCCCTTGGAGG 1191
 QY 1195 ATGTACATGAATCATCCATCTGATGACAGAGGAGAGACCACT 1241
 Db 1192 CTTTATAATAAATTTAATCTAATGAGGAATATACAGTGAAGCT 1238

RESULT 7

HLXYACLY
 LOCUS
 DEFINITION H.sapiens (HepG2) LAL mRNA for lysosomal acid lipase.
 ACCESSION Z31690
 VERSION Z31690.1 GI:506430
 KEYWORDS LAL; lipase; lysosomal acid lipase.
 SOURCE Homo sapiens (human)

ORGANISM
 Homo sapiens

REFERENCE
 1. (bases 1 to 2481)

AUTHORS
 Du, H. and Gregory, G. A.

TITLE
 Structural Conservation of Putative Functional Motifs between Mouse and Human Lysosomal Acid Lipase

JOURNAL
 Unpublished

REFERENCE
 2. (bases 1 to 2481)
 Du, H.
 Direct Submission
 Submitted (05-APR-1994) Hong Du, Division of Human Genetics,
 JOURNAL

HLXYACLY 2481 bp mRNA linear PRI 17-FEB-1997
 LOCUS H.sapiens (HepG2) LAL mRNA for lysosomal acid lipase.

DEFINITION H.sapiens (HepG2) LAL mRNA for lysosomal acid lipase.
 ACCESSION Z31690
 VERSION Z31690.1 GI:506430
 KEYWORDS LAL; lipase; lysosomal acid lipase.
 SOURCE Homo sapiens (human)

ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 1. (bases 1 to 2481)
 Du, H. and Gregory, G. A.
 Structural Conservation of Putative Functional Motifs between Mouse and Human Lysosomal Acid Lipase

JOURNAL
 Unpublished
 REFERENCE
 2. (bases 1 to 2481)
 Du, H.
 Direct Submission
 Submitted (05-APR-1994) Hong Du, Division of Human Genetics,
 JOURNAL

Children's Hospital Medical Center, 3333 Bernet Street, Cincinnati,
OH, 45229-3039, USA

FEATURES

Location/Qualifiers
1. .2481
/organism="Homo sapiens"
/mol_type="mRNA"
/strain="HepG2"
/db_xref="taxon:9606"
/clone="clone hLAL"
/cell_type="hepatoma"
/tissue_type="liver"
/clone_lib="human liver HepG2 cDNA library/lambda ZAP"
1. .2481
/gene="LAL"
34. .1233
/gene="LAL"
/citation=[1]
/codon_start=1
/product="lysosomal acid lipase"
/protein_id="CAA83495.1"
/db_xref="GI:506431"
/db_xref="GPI:016529"
/translation="MKMFLGLVCLVLTLSHSGSKLTAVDPETNMVSEIISVM
GPFSEYLVEDGVLCLNRIPIGRKNSDKPKPVVFLQHLGLASSNWTNLS
SIGFLADAGEDVMNGSRNTWSRKHKLTVSQDEFWAFSYDMKADIDLPASINFL
NKTGEQYVYVHSGQTTFGTFAPSLQELAKRIMFALGFVASFVAFSPMAKGR
LPDLIKDLFGDKPELPQSAFLKMLGHVTHVILKELGNLFCFLCGFNERNLNMR
VDVYTHSPAGTSVQNMHLSQAVKFKQFAPDMGSSAKNFYHNSQVPPYNYKML
VPTAVSGSHDLADVVDVNIQLTQITNLVPHESIPEWHLDFIWLDPRLNKKII
NLNRKYQ"
34. .96
/gene="LAL"
/citation=[1]
97. .1230
/gene="LAL"
/product="lysosomal acid lipase"
/citation=[1]
79
/gene="LAL"
/citation=[1]
/replaces="g"
2459. .2464
/genes="LAL"
2480
/gene="LAL"
692 a 477 c 513 g 799 t

sig_peptide

mat_peptide

variation

variation

polyA_signal

polyA_site

BASE COUNT
ORIGIN

Query Match 41.6%; Score 527.8; DB 9; Length 2481;
Best Local Similarity 66.6%; Pred. No. 1e-151;
Matches 751; Conservative 0; Mismatches 376; Indels 0; Gaps 0;
Qy 115 AAAGCTGTGACCCAGAGCATTCATGAATATTAGTGAATCATCCAAATCAAGCTAT 174
Db 112 ACAGCTGTGGATCCTGAAACAAATGATGAGTGAATATCTCTTACTGGGATTC 171
Qy 175 CCTGTGAGGAATATGAAGTCGCAATGAAGATGGGTATATCTCTTAAACAGGATT 234
Db 172 CCTAGTGAAGATACCTAGTTGAGACAGAGATGATATATCTGTGCCTTAAACGGAAT 231
Qy 235 CCTCGAGGCTAGTGACCACTAAGACAGAGTTCCAGGCTGTGTGTTTACTTGCAGCAT 294
Db 232 CCTATCGGAGGAGAACCAATCTGACAAAGTCCCAACACAGTTGTCTTCTGCAAT 291
Qy 295 GGCTAGTTGGAGTGTAGCAATCTGCAATCTGCAATCTGCAATCTGCAATCTGCAAT 354
Db 292 GGCTTGTGCGAGATTCTAGTAAGTGGTCAAAAACCTTGCACACAGAGCTGGGCTTC 351
Qy 355 ATTCTGCAGATGCTGCTTTTGGAGTGGATGGGACAGAGGGGAAACGCTGTCT 414

Db 352 ATCTGTGTGATGCTGCTGCTTTTTCACCTGTGTGGTGGCAACAGCAGAGGAAATACCTGGTCT 411
Qy 415 CGAAACACAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTCAGTTATCATGAG 474
Db 412 CGGAAACATAAGACACTCTCAGTTTCTCAGGATGAATTTCTGGGCTTTCAGTTATCATGAG 471
Qy 475 ATGGCTAGGTTTGAACCTTCTCAGTGTATTAATTTATTTGACAGAAACGGCCAGAA 534
Db 472 ATGGCAAAATATGACCTACAGCTTCCATTAATTTCAATTTCAATTAATTTGCAAAATG 531
Qy 535 AGATCTATATGTGGCTTATTCAGGSCACCAACATGGGCTTTTATGCAATTTCCACC 594
Db 532 CAAGTGTATATGTGGTCAATTTCAAGGCAACCATAGGTTTATAGCATTTTCAACAG 591
Qy 595 ATGCCAGAGCTGGCTCAGAAATCAAAATGTATTTTGTCTTATAGCACCCATAGCACTGTT 654
Db 592 ATCCCTGAGCTGGCTAAAGAGATTAATTTTGTCTTGTGGCTTCTGTGGCTTCCGTC 651
Qy 655 AAGCATGCAAAAGCCCGGACCAAAATTTTGTCTGCTGCCAGATATGATGATCAAGGA 714
Db 652 GCCTTCTGTACTAGCTTATGGCCAAATTTAGGACGATTAACAGATCATCTCATTAGGAC 711
Qy 715 TTCTTTGGCAAAAGAAATTTCTGTATCAGACCAAGATTTCTCAGACAACTTGTATTATAC 774
Db 712 TTATTTGGACAAAGAAATTTCTCCAGAGTGGCTTTTGAAGTGGTGGTACCCAC 771
Qy 775 CTTTGTGGCCAGGTGATTTCTGATCAGATTTGTAGTAATATCATGTTATCTTCTGGTGA 834
Db 772 GTTTGCATCATCTCATCTGAAGGAGCTCTGTGGAATATCTGTGTTTCTTCTGTGGA 831
Qy 835 TTCAACACCAATATGATGACATGAGCCGAGCAAGTATATGCTGCCACACACTCTTGTCT 894
Db 832 TTTAATGAGAGAAATTTAAATATGCTAGATGGATGATATACACATTTCTCTCTGCT 891
Qy 895 GGAACATCTGTGCAAAATATTTACACTGAGCAGGAGTGAATTTCTGGTGAATCTCCGG 954
Db 892 GGAATCTGTGCAAAACATGTTTACACTGAGCAGGCTGTAAATTTCCAAAAGTTTCAA 951
Qy 955 GCATTTGACTGGGGAGTGAGACCAAAATCTGAAAATGCAATCAGCCAACTCCCTGA 1014
Db 952 GCCTTTGACTGGGGAAGCAGTGCCAGAAATTTATTTTATTACCAACAGAGTTATCTCCC 1011
Qy 1015 AGGTACAGAGTCAGATATGAGGTCCTTACACCAATTTGGAAGAGGTCAGAGCTGG 1074
Db 1012 ACATACATGTGAGGACATGCTTGTGCGGACTGCACTCTGAGCGGGGTCACGACTGG 1071
Qy 1075 CTTTCAAAATCCAGAGACGTGAAAATGCTGCTCTGAGGTGACCAACCTCATCTACCAT 1134
Db 1072 CTTGAGATGCTTACGAGCTCAATATCTTACTGACTCAGATCACCACCTTGGTGTCCAT 1131
Qy 1135 AAGAAATTTCTGAATGGGCTCAGTGGATTTTCACTGGGGTTTGGATGCTCTCCACCGT 1194
Db 1132 GAGAGCAATTCGGAAATGGGAGCATCTTGAATTTATTTGGGCTTGGATGCCCTTGGAGG 1191
Qy 1195 ATGTACAAATGAATTCATTCATCTGATGTCAGCAGAGGAGACCAACCT 1241
Db 1192 CTTTATAATAAAATTAATTAATTAATGAGGAAATATCAGTGAAGCT 1238

RESULT 8

HSU08464
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

HSU08464
Human lysosomal acid lipase mRNA, complete cds.
U08464
U08464.1 GI:505052
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2481)
Du, H. and Gregory, G.A.

HSU08464
2481 bp
mRNA
linear
PRI 23-JUN-1994
complete cds.

AUTHORS Amais,D.
 TITLE Direct Submission
 JOURNAL Submitted (29-NOV-1993) D. Amais, Medical Department, University
 Hospital Eppendorf, Martinistrasse 52, 20246 Hamburg, FRG

FEATURES
 source
 1..2626
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="10"
 /tissue_type="liver"
 1..2626
 /gene="LAL/LIPA"
 146..1345
 /gene="LAL/LIPA"
 /EC_number="3.1.1.13"
 /note="sterol esterase"
 /codon_start=1
 /evidence=experimental
 /product="lysosomal acid lipase"
 /protein_id="CAA54026.1"
 /db_xref="GI:434306"
 /translation="MKRFLGLVCLVLTLSHSGSKLTAVDPETNMVSEIISW
 GPSEYIVETEDGYILCLNRI PHGRKNHSDKPKVPVFLQHGGLADSSNWNILANS
 SLGFLADAGFDVWGRSGRNTWSRKHKLVSQDEFWA PSYDEMAYDILPASINFIL
 NKTGQEQYVGHSGQTTIGTIFASQIPFLAKRIMFALGSPVAVAFCTSPRAKGR
 LPDHLIKDLFGDKBFLPOSALFKWLGVTHVTHILKELCNLCFLCGFNERNLMSR
 VDVYTHSPAGTSVQNLHWSQAVKFKQFQADWMSAKNYFHYNOSYPTTYNPKML
 VPTAVSGGHDWLDVYDYNILLTQITNLVPHESIPEWEHLDFIWLGLDAPWRLYNKII
 NLMRYQ"
 146..226
 /gene="LAL/LIPA"
 /evidence=experimental
 374..1342
 /gene="LAL/LIPA"
 /product="lysosomal acid lipase"
 /evidence=experimental
 227..373
 /gene="LAL/LIPA"
 /note="propeptide"
 /evidence=experimental
 741 a 527 c 547 g 811 t

BASE COUNT 741 a 527 c 547 g 811 t
 ORIGIN

Query Match 41.6%; Score 527.8; DB 9; Length 2626;
 Best Local Similarity 66.6%; Pred. No. 1e-151;
 Matches 751; Conservative 0; Mismatches 376; Indels 0; Gaps 0;

QY 115 AAAGCTGTGGACCCAGAGCAATTCATGAATATTAGTGAATCAATCAACATCAAGGCTAT 174
 DB 224 ACAGCTGTGGATCTCTGAAACAAACATGAATGTGAGTGAATATCTCTTACTGGGGATTC 283

QY 175 CCTGTGAGGATATGATGCTCCCACTGAGATGGGTATATCCCTTCTGTAAACAGGATT 234
 DB 284 CTTAGTGAAGATACCTAGTTGAGACAGAGATGATATATCTGTGCTTAAACCGAATT 343

QY 235 CCTCAGGCGCTAGTCAACCTTAAGAACAGACAGGTTCCAGGCGCTGTGGTGTACTGCAGCAT 294
 DB 344 CCTCATGGCAGGAGAACCACTCTGCACAAAGGTCCCAACCAAGTTGTCTTCTCTGCAACAT 403

QY 295 GGCCTAGTTGGAGGTGCTAGCAACTGGATTTCACACCTGCCCAACATAGCTGGGCTTC 354
 DB 404 GCGTGTCTGGCAGATCTTAGTAACTGGGTGCACAAACCTTGCCCAACAGCAGCTGGGCTTC 463

QY 355 ATTCTGGCAGATGCTGGTGTGACCTGTGGATGGGGAACAGCAGGGGAAAGCGCTGTCT 414
 DB 464 ATCTCTGCTGATGCTGGTGTGACCTGTGGATGGGGAACAGCAGAGGAATACCTGTCT 523

QY 415 CGAAACACACAGACACTCTCCATAGACACAGATGATGTTCTGGGCTTTCAGTTATGATGAG 474
 DB 524 CGAAACACATAGACACTCTCAGTTCTCAGGATGAATCTGGGCTTTCAGTTATGATGAG 583

QY 475 ATGGCTAGTGTGACCTCTCTCAGTATGAATACTTTATTTTGCAGAAACGGCCAGGAA 534
 DB 584 ATGGCAAAATATGACCTACAGCTTCCATTAACTTCAITCTGAATAAAATGGCCAGAA 643

QY 535 AAGATCTATATGTGGCTATTCACAGGCAACCACTGGGCTTTATTTGCATTTTCCACC 594
 DB 644 CAAGTGATATATGTGGGCTCAITCTCAAGGCACCACTAGGTTTTATAGCATTTTCAAG 703

QY 595 ATGCCAGAGCTGGCTCAGAAAATCAAAATGTATTTTGTGTTTAGCACCCTATGCCACTGTT 654
 DB 704 ATCCCTGAGCTGGCTAAAGGATTAATAATGTTTTTGGCCCTGGGCTGTGGCTTCCGTC 763

QY 655 AAGCATGCAAAAGCCCGGACCAAAATTTTGTGTTGCTGCCAGATATGATGATCAAGGA 714
 DB 764 GCTTCTGTACTAGCCCTATGCCAAATTTAGCAAGATTTACCAAGATCATCTCATTAGGAC 823

QY 715 TTGTTGGCAAAAAGAAATTTCTGTATCAGACCAAGATTTCTCAGACAACTTCTTTATTAC 774
 DB 824 TTATTTGGAGACAAAGAAATTTCTTCCCCAGAGTGGCTTTTGAAGTGGCTGGTACCCAC 883

QY 775 CTTTGGCCAGGTGATTTCTTGATCAGATTTGTAGTAAATATCATGTTTACTTCTGGTGA 834
 DB 884 GTTTGCACTCATGTCTACTGAAGGAGCTCTGTGAAATCTCTGTTTTCTTCTGTGTGA 943

QY 835 TTCAACACCAACAATATGAATGAATGAGCCGAGCAAGTGTATATGCTGCCACACTCTTGCT 894
 DB 944 TTTATGAGAGAAATTTAAATATGTCTAGAGTGGATGATATACAACATTTCTCTGCT 1003

QY 895 GGAACATCTGTGCAAAATATTTCACTGAGACCCAGGAGTGAATTTGTGGTGAATCCGG 954
 DB 1004 GGAATCTCTGTGCAAAACATGTTTACACTGGAGCCAGGCTGTTAAATTTCCAAAAGTTTCAA 1063

QY 955 GCATTTGCTGGGGAGTGAACCAAAATCTGGAATAATCAATCAGCAACTCTCTGTA 1014
 DB 1064 GCTTTGACTGGGGAAGCAGTCCCAAGAAATTAATTTTCAACACAGATTAATCTCTCC 1123

QY 1015 AGGTACAGAGTCAAGATATGACGCTCCCTCAGCAATGTGACAGAGGAGTCAAGACTGG 1074
 DB 1124 ACATCAATGTGAAGGACATGCTTGTGCCGACTGCGAGTCTGGAGGGGGGTCAAGACTGG 1183

QY 1075 CTTTCAATCCAGAGACGTGAAATGCTCTCTGAGGTGACCAACCTCATCTACCAT 1134
 DB 1184 CTTGCAAGTGTCTACGAGCTCAATATCTTACTGACTCAGATCACCACCTTGTGTCTCAT 1243

QY 1135 AAGATATTTCTGAATGGGCTCACGTGGATTTTCACTGGGGTTTGGATGCTCTCACCGT 1194
 DB 1244 GAGAGCATTCGGATGGGAGCATCTTGACTTCAITTTGGGGCTGATGCCCTTGGAGG 1303

QY 1195 ATGTCAATGAATATCATCTCTGATGACGAGGAGGAGACCAACT 1241
 DB 1304 CTTTATAATAAATTAATTAATCTAATGAGGAAATATCAGTGAAGCT 1350

RESULT 10
 BD094072 2493 bp DNA linear PAT 27-AUG-2002
 LOCUS Shear stress-responsive DNAs.
 BD094072
 ACCESSION BD094072.1 GI:22639660
 VERSION WO 0125427-A/33.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 2493)
 Nojima,H., Yoshisue,H., Obayashi,M., Ota,T., Kawabata,A.,
 Sakurada,K., Kuga,T., Sekine,S., Nakamura,Y. and Sugan,S.
 Shear stress-responsive DNAs
 Patent: WO 0125427-A 33 12-APR-2001;
 KYOWA HAKKO KOSYO CO LTD,HIROSHI NOJIMA,HAJIME YOSHISUE, MASAYA
 ODAYASHI, TOSHIO OTA,AYAKO KAWABATA,KAZUHIRO SAKURADA,TETSURO KUGA,
 SUSUNO SEKINE, YUSUKE NAKAMURA,SUMIO SUGANO
 OS Homo sapiens (human)

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

PN WO 0125427-A/33
 PD 12-APR-2001
 PF 02-OCT-2000 WO 2000JP006840
 PR 01-OCT-1999 JP 99P 280976
 PI HIROSHI NOJIMA, HAJIME YOSHISUE, MASAYA OBAYASHI, TOSHIO OTA, PI
 AYAKO KAWABATA,
 PI KAZUHIRO SAKURADA, TETSURO KUGA, SUSUMU SEKINE, YUSUKE NAKAMURA,
 PI SUMIO SUGANO
 PC C12N15/12, C07K14/435, C07K16/18, C12P21/02, C12Q1/69, A61K38/00,
 PC A61K39/395,
 PC A61K48/00, A61P9/10, G01N33/50, G01N33/53
 CC
 FH
 FT

Key Location/Qualifiers
 CDS Location/Qualifiers
 1. .2493
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 692 a 484 c 516 g 801 t

BASE COUNT
 ORIGIN

Query Match 41.5%; Score 526.2; DB 6; Length 2493;
 Best Local Similarity 66.8%; Pred. No. 3.1e-151;
 Matches 751; Conservative 0; Mismatches 376; Indels 0; Gaps 0;

QY 115 AAGCTGTGACCCAGCAATTCATGAATATTAGTGAATATCCAAATCAAGGCTAT 174
 DB 119 ACAGCTGTGATCTCTGAAACAAACATGAATGTGAGTGAATATCTCTTACTGGGATTC 178

QY 175 CCCTGTGAGCAATGAAGTCGCACTGAAGATGGGTATATCTCTCTTAACAGGATT 234
 DB 179 CCTAGTGAGCAATCTAGTTGAGACAGAGATGATATCTGTGCGCTTAACGAATT 238

QY 235 CCTCGAGGCTAGTGCACCTTAAGACAGAGAGTTCAGGCGCTGTGTGTATCTGACGAT 294
 DB 239 CCTCATGGGAGGAAGAACCTTCTGACAAAGTCCCAACAGTGTCTCTCTCTGCAACAT 298

QY 295 GGCTAGTGTGAGTGTCTGACCACTGATTTCCAACTGCTCCCAACATAGCTGGGCTTC 354
 DB 299 GGCTGTGCGAGATCTAGTAACTGGGTCAAAACCTTCCCAACAGAGCTGGGCTTC 358

QY 355 ATCTGCGCAGATGCTGTTTTGACGTGTGATGGGAAACAGAGGGGAAACGCTGGTCT 414
 DB 359 ATCTGCTGATGCTGTTTTGACGTGTGATGGGAAACAGAGAGGAAATACCTGGTCT 418

QY 415 CGAAACACAGACACTCTCCATGACCAAGATGAGTCTGGGCTTTCAGTTATGATGAG 474
 DB 419 CGAAACACATAGACACTCTCAGTTTCTCAGGATGAATCTGGGCTTTCAGTTATGATGAG 478

QY 475 ATGCTAGGTTTGACCTTCTGCGAGTGAATAAATTTTATTTTGCAGAAACGGGCCAGGAA 534
 DB 479 ATGCGAAATATGACCTACCGATTCCTCACTTAATCTTCACTGAATAAATCTGGCCAGAA 538

QY 535 AAGATCTATATGTCGGCTATTCAAGGCAACACAGGCTTATGCAATTTTCCACC 594
 DB 539 CAAGTGTATTATGTGGGTCTATTCTCAAGGCAACCACTATAGGTTTTATAGCATTTTCAAG 598

QY 595 ATGCGAGAGTGGTTCAGAAATCAAAATGATTTTGTCTTTAGCCCATAGCACTGTT 654
 DB 599 ATCTGAGCTGGCTTAAAGATTAATGTTTTTTCCTGGTCTGGTCTGGCTCCGTC 658

QY 655 AAGCATGCAAAAGCCCGGACCAAAATTTTGTGTGTCGAGATATGATGATCAAGGGA 714
 DB 659 GCCTTCTGCTAGTCCCTATGGCCAAATTTAGGAGGATTAACAGATCATCTCATTAAGGAC 718

QY 715 TTGTTTGCAGAAAGATTTCTGATCAGACAGATTTCTCAGACAACTGTTATTATTC 774
 DB 719 TTATTTGAGCAAGAAATTTCTCCCGAGAGTGGTTTTTGAAGTGGCTGGTACCCAC 778

QY 775 CTTTGTGCGCAGGATCTTGTATCAGATTTGTATGATATCAATGTTACTTCTGGTGA 834
 DB 779 GTTTGCACTCATGCTACTGAAGAGCTCTGTGAAATCTCTGTTTTCTTCTGTGTGGA 838

QY 835 TTCAACACCAACAATATGAACATGAGCCGAGCAAGTGTATATCTGCTCCACACTCTTGTCT 894
 DB 839 TTTAATGAGAGAAATTTAATATATGCTAGAGTGTATATACCAACATCTCTCTGCT 898

QY 895 GGAACATCTGTGCAAAATATTCTACACTGGAGCCAGGAGTGAATTTCTGGTGAACCTCGG 954
 DB 899 GGAACATCTGTGCAAAACATGTTTACACTGGAGCCAGGCTGTTAAATTTCCAAAGTTTCAA 958

QY 955 GCATTTGACTGGGGAGTGAACCAAAATCTGAAAAATGCAATCAGCAACACTCTCTGTA 1014
 DB 959 GCCTTTGACTGGGAGGAGCAGTGCACAGAAATTTTTCATTAACACAGAGTTATCTCTCC 1018

QY 1015 AGGTACAGATGACAGATATGACGGTCCCTACAGCAATGTGAGCAGAGGTGAGGACTGG 1074
 DB 1019 ACATACAAATGTGAAGGACATGCTTGTGCGGACTGCACTCTGGAGCGGGGTACGACTGG 1078

QY 1075 CTTTCAATCCAGAGACGTGAAAAATGCTCTCTGAGGTGACCAACCTCATCTCACTCAT 1134
 DB 1079 CTTGCAATGCTTACGACGTCAATATCTTACTGACTCAGATCACCACCTTGGTGTCCAT 1138

QY 1135 AAGATATCTCTGAATGGGCTCAGTGGATTTCACTGGGGTTTGGATGCTCTCACTGCT 1194
 DB 1139 GAGAGCATCTCGGAATGGGAGCATCTTCACTTCTGCGGCTGGATGCTGCTGGAGG 1198

QY 1195 ATGTACATGAATCACTCCATCTGATGACGAGGAGGAGACCAACT 1241
 DB 1199 CTTTATTAATAATTTAATCTAATGAGGAATATCAGTGAAGCT 1245

RESULT 11
 HUMLIPLCHL 2493 bp mRNA linear PRI 07-JAN-1995
 LOCUS Human lysosomal acid lipase/cholesterol ester mRNA, complete
 DEFINITION cds.
 ACCESSION M74775
 VERSION M74775.1 GI:187151
 KEYWORDS lysosomal acid lipase/cholesterol esterase.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (Bases 1 to 2493)
 AUTHORS Anderson, R.A. and Sando, G.N.
 TITLE Cloning and expression of cDNA encoding human lysosomal acid
 lipase/cholesterol ester hydrolase. Similarities to gastric and
 lingual lipases
 JOURNAL J Biol. Chem. 266 (33), 22479-22484 (1991)
 MEDLINE 92042392
 PUBMED 1718995
 COMMENT Original
 FEATURES
 source
 1. .2493
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /map="10"
 /tissue_type="fibroblast"
 1. .2493
 /gene="LIPA"
 <1. .2493
 /gene="LIPA"
 /note="G00-120-153"
 41. .1240
 /gene="LIPA"
 /EC number="3.1.1.13"
 /note="alternate initiator ATG may be at position 47"
 /codon_start=1
 /product="lysosomal acid lipase/cholesterol esterase"
 /protein_id="AA59519.1"
 /db_xref="GI:187152"
 /db_xref="GDB:G00-120-153"
 /translation="MKRFLGLVCLVLPVHSEGGGKLTAVDPETNMNVSELIISW"

115	AAAGCTGTGAGCCAGGAAGCATTCATGAAATATTTAGTGAATCATCCAAACATCAAGGCTAT	174
Qy		
119	ACAGCTGTGATCTCTGAAACAAACATGAATGTGAGTGAATATCTCTTACTGGGGAATC	178
Db		
175	CCCTGTGAGGAATATGAAGTCGCAACTCAAGATGGGTATATCCCTTTCTGTTTAAACAGGATT	234
Qy		
179	CCTAGTCAGGAATACCTTAGTTTGAGACAGAAGATGATATATCTGTGCCTTAAACCGAATT	238
Db		
235	CCTCGAGGCCTTAGTGCAACCTAAGAAAGACAGTTCCAGGCGCTGGTGTTACTTCGCAGCAT	294
Qy		
239	CCTCATGGGAGGAAGAACCATTCGTGACAAAGGTCCTCAACACAGTGTCTCTCTGCAACAT	298
Db		
295	GGCCTAGTTGGAGTGCTAGCAACTGGATTTCCAACTGCCCAACAATAGCCTGGGCTTC	354
Qy		
299	GGCTTGCTGGCAGATTCTAGTAACTGGGTCAAAACCTTGCCAAACAGACGCTGGGCTTC	358
Db		
355	ATTCTGCGAGATGCTGGTGTTTGACGTGTGGATGGGAACAGCAGGGGAAAGACCTCGTCT	414
Qy		
359	ATTCTTGCTGATGCTGGTGTTTGACGTGTGGATGGGCAACAGCAGAGAAATACCTGGTCT	418
Db		
415	CGAAAAACAAGACACATCTCCATAGACCAAGATGAGTCTGGGCTTTTCAGTTATGATGAG	474
Qy		
419	CGAAACATAAGACACATCTCAGTTTCTCAGGATGAATCTGGGCTTTTCAGTTATGATGAG	478
Db		
475	ATGGCTAGTTTGACCTTCTCGCAGTGATAAATTTTATTTTGCAGAAAAACGGGCGACGAA	534
Qy		
479	ATGGCAAAATATGACCTACCAAGCTTCCATTAACTTCACTCTGATAAAACCTGGCCAGAA	538
Db		
535	AAGATCTATTATGTGGGCTATTTCACAGGGCACCAACCTATGGGCTTTATGCAATTTCCACC	594
Qy		
539	CAAGTGATATTATGTGGGTCAATCTCACAAGGCCACCACTATAGGTTTATAGCAATTTTTCAG	598
Db		
595	ATGCCAGAGCTGGGTTCAGAAAATCAAAATGTATTTTGGCTTTTTCAGCACCCATAGCCACTGTT	654
Qy		
599	ATCCCTGAGCTGGCTAAAGGATTAATAATGTTTTTGGCCCTGGGTCTGTGGCTTCCGTC	658
Db		
655	AAGCATGCAAAAGACGCCCGGACCAAAATTTTGTGTGCTGCCAGATATGATGATCAAGGGA	714
Qy		
659	GCCTTCTGTACTAGCCCTATGGCCAAATTAGGACGATTACAGATCATCTCATTTAAGAC	718
Db		
715	TTGTTTGGCAAAAAAGAAATTTCTGTATCAGACCAAGATTTCTCAGACAACTTGTATTATTAAC	774
Qy		
719	TTATTTGGAGACAAGAAATTTCTTCCCAAGAGTGGTGTTTTGAAGTGGCTGGGTACCCAC	778
Db		

passed the following selection criteria: matched mRNA gi: 434305.

FEATURES
source
Location/Qualifiers
1..2586
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:5229 IMAGE:2900168"
/tissue_type="Placenta, choriocarcinoma"
/clone_id="NIH MGC_10"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
101..1300
/codon_start=1
/product="Similar to lipase A, lysosomal acid, cholesterol
esterase (Simlan disease)"
/protein_id="AAH2287.1"
/db_xref="GI:15126727"
/translation="MKRPLGLVCLVLPVHSEGGGKLTALDPETNMNVSIIISYW
GFPSEYLVETEDYIILNRIPIRKHSHDKPKPVVFLQHLGLADSNWVNTLANS
SLGFIADAGDVHMSGRNTRKXKHTLSVQDEFWAFSDYEMAKYDLPASINFL
NKTGOVYVGHSGQTTIGFAPQIPELAKRIKMPFALGPVAVAFCTSPMAKLR
LPDLIKDLKDFELPQSAFLKWLGVTHVTLKELCNLCFLCGLFERNLNMR
VDVYTHSPAGTSVQNLHWSQAVKQFQAFDWSAKNYPHYNOSYPTVNVKML
VPTAVSGSDHDLADVDVNIILQIINLVFHSIPEWEHLDFINGLDAPWRLNKLII
NLMKRYQ"

CDS
BASE COUNT 736 a 511 c 531 g 808 t
Query Match 41.3%; Score 524.6; DB 9; Length 2586;
Best Local Similarity 66.5%; Pred. No. 9.9e-151;
Matches 750; Conservative 0; Mismatches 377; Indels 0; Gaps 0;

115 AAAGTGTGGAGCCAGAGCATTCATGAATATTAGTGAATCATCCAAATCAAGGCTAT 174
179 ACAGCTTTGGATCTCTGAAACAAACATGAATGAGTGAATATCTCTTACTGGGATTC 238
175 CCCTGTGAGGATATGAGTGCACACTGCAAGTGGTATATCTCTTCTGTAAACAGATT 234
239 CTTAGTGGAGGATACCTAGTTGAGACAGAGATGATATATCTGTGCTTAAACGGAAT 298
235 CTTGAGGCTAGTGCACACTGAAAGACAGGTTCCAGGCTGTGGTGTACTGACGAT 294
299 CTTGATGGGAGGAGAACCATCTTGACAAAGGTCCCAACCCAGTTGCTCTCTGCAACAT 358
295 GGCCTAGTTGAGGTTGATGCACTGATTTCCACCTGCTCCCAACATAGCTGGGCTTC 354
359 GGCCTGCTGGCAGATTTAGTAACCTGGGTCAAAACCTTGCCACAGAGAGCTGGGCTTC 418
355 ATTCTGGCAGATGCTGGTGTGTTGACGTGTGGATGGGAAACAGCAGGGAACCCCTGGTCT 414
419 ATTCTGTGATGCTGGTGTGTTGACGTGTGGATGGGAAACAGCAGGGAATACCTGGTCT 478
415 CGAAACACAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTCAGTTATGATGAG 474
479 CGAAACACAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTCAGTTATGATGAG 538
475 ATGGCTAGGTTGACCTTCTCAGTGTAACTTTATTTTTCAGAAACGGGCGAGAA 534
539 ATGGCAATATGACCTACAGCTTCCATTAATCTTCACTTCAATTAATTAATGGGCAAGAA 598
535 AAGATCTATATGCTGGCTTATTCAGGGGCAACCATGAGGCTTTTATGATTTTCCACC 594
599 CAAGTGATATATGCTGGCTTATTCAGGGGCAACCATGAGGCTTTTATGATTTTCCACC 658
595 ATGCCAGAGCTGGCTCAGAAATCAAAATGATATTTTGTGTAGCACCCTAGCCACTGTT 654
659 ATCCCTGAGCTGGCTAAAGGATTAATATGTTTTTGGCCCTGGGTCCTGGCTTCCGTC 718
655 AAGCATGCAAAAGCCCGGAGCCAAATTTTGTGCTGCCAGATATGATGATCAAGGGA 714
719 GCCTTCTGACTAGCCCTATGCCAAATTAGGACGATTACAGATCATCTCATTAAGGAC 778
715 TTGTTGGCAAAAGAAATTTCTGATATCAGACCAAGTTCTCAGACCAACTTGTATTAC 774

Db 779 TTAATTGGAGACAAGAAATTTCTTCCCAGAGTGGTGTGAGTGGGTACCCAC 838
QY 775 CTTTGTGCCAGGTGATCTTGTATCAGATTGTGTAGTAAATATCATGTTACTCTGGTGA 834
Db 839 GTTTCACATCTATGCTACTGAGGAGCTCTGTGAAATCTCTGTTTCTTCTGTGGA 898
QY 835 TTCAACACCAACAATATGAACATGAGCGGAGCAAGTGTATATGCTGCCACACATCTTGTCT 894
Db 899 TTTAATGAGAGAAATTTAAATATGTCTAGAGTGGATGTATATACACACATTTCTCTGCT 958
QY 895 GGAACATCTGTGCAAAATATTCTACATGAGGAGCAGGAGTGAATTTCTGGTGAATCTCGG 954
Db 959 GGAACATCTGTGCAAAATATTCTACATGAGGAGCAGGAGTGAATTTCTGGTGAATCTCGG 1018
QY 955 GCATTTGACTGGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCCAACTCTCTGTA 1014
Db 1019 GCCTTTGACTGGGAGAGCAGTCCCAAGAAATTTTCTTATACACAGAGTATCTCTCCC 1078
QY 1015 AGGTACAGAGTCAGAGATATGACGGTCCCTACAGCAATGTGGACAGGAGTCAAGACTGG 1074
Db 1079 ACATACAATGTGAAGGACATGCTTGTGCCGACTGCAAGTCTGGAGCGGGGTCAAGACTGG 1138
QY 1075 CTTTCAATCCAGAGAGCTGAAATGCTCTCTGAGGTGACCAACCTCATCTACCAT 1134
Db 1139 CTTGAGATGCTACAGCTCAATATCTTACTGACTCAGATCAACACTTGTGTTCAT 1198
QY 1135 AAGAATATTCTGAATGGGCTCACTGATTTCACTGGGTTTGGATGCTCTCCACCGT 1194
Db 1199 GAGAGCATTCGGATGGGAGCATCTTGACTTCTATTTGGGCGCTGGATGCCCTTGGAGG 1258
QY 1195 ATGTACAATGAATCATCTCATCTGATGAGGAGGAGACCAACCT 1241
Db 1259 CTTTATAATAAATTAATTAATCTAATGAGGAATATCAGTGAAGCT 1305

RESULT 13

LOCUS A26689 1378 bp DNA linear
DEFINITION Precursor of rabbit gastric lipase coding sequence.
ACCESSION A26689
VERSION A26689.1 GI:905029
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM Benicourt, C., Blanchard, C. and Junien, J.L.
REFERENCE 1 (bases 1 to 1378)
AUTHORS Recombinant gastric lipase from rabbit and pharmaceutical
TITLE Patent: EP 0542629-A 9 19-MAY-1993;
JOURNAL INSTITUT DE RECHERCHE JOUVEINAL
FEATURES
source
1..1378
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
BASE COUNT 397 a 276 g 418 t
ORIGIN

Query Match 38.6%; Score 489.2; DB 6; Length 1378;
Best Local Similarity 65.0%; Pred. No. 8.6e-140;
Matches 722; Conservative 0; Mismatches 388; Indels 0; Gaps 0;
QY 125 ACCGAGAGCATTCATGAATATTAGTGAATCATCCAAATCAAGGCTATCTCTGTGAGG 184
Db 136 ACCCTGAAGTGAATGAATATTAGTCAAGTATTTCTTCTGAGGATACCAAGTGA 195
QY 185 AATATGAGTGCACACTGAAGTGGTATATCTTCTGTTTAAAGGATTTCTCGAGCC 244
Db 196 AATATGAGTGTGACTGAAGTGGTATATCTTGAAGTAAACAGAAATTTCTTATGGA 255
QY 245 TAGTGCAACCTAAGAAGACAGGTTCCAGGCTGTGGTGTACTGACGATGCGCTAGTTG 304

Db 256 AGAAAAATTCAGGAAACAGAGCCAGAGACCGTGTATTTTTCAGCATGTTGCTTG 315
 Qy 305 GAGGTCTAGCACTGGATTTCCAACTGCCCCAAATAGAGCTGGCTTCATTCGGCAG 364
 Db 316 CATGACCTCAAACTGGATCTCCAACTGCCCCAAACAGAGCTGGCTTCATTCGGCAG 375
 Qy 365 ATGCTGTTTTCAGCTGTGGATGGGAAACAGAGGAGAAACCCCTGGTCTGAAACACA 424
 Db 376 ATGCTGTTTTCAGCTGTGGATGGGAAACAGAGGAGAAATACCTGGTCCAGAGAACT 435
 Qy 425 AGACACTCTCCATAGACCAAGATGAGTTCTGGCTTTTCAGTTATGATGAGATGGTAGGT 484
 Db 436 TATACATTCACAGACTCCGTTGAATTCCTGGCTTTTAGCTTTTCATGAAATGGCTAAAT 495
 Qy 485 TTGACCTTCTGTCAGTGAATAAATTTTATTTTCAGAAACCGGCGAGGAAAGATCTATT 544
 Db 496 ATGACCTTCCAGCCCAATGACTTCATTGTAAGGAACTGGACAGGAAAGCTGCACT 555
 Qy 545 ATGTCGGCTATTACAGGAGCACCACCATGGCTTTTATTCATTTTCCACCATGCCAGAGC 604
 Db 556 ATGTTGGTCAATCCAGGAGCACCACCATGGCTTTTATTCCTTTTCTACAAATCCCAAGC 615
 Qy 605 TGCTCAGAAATCAAAATGATTTTTCCTTTAGCAGCCCATAGCCTCTGTAAGCATGAA 664
 Db 616 TGCTGAAAGATCAAAATGATTTTTCCTTTAGCAATGACTCCAGTCCCACTGTGAATACAA 675
 Qy 665 AAGCCCGGAGCAAAATTTTGTGCTGCCAGATATGATGATCAAGGATGTTTGGCA 724
 Db 676 AAGCCCTGTGAAACAAATAGCTTTTATTCCTCCAACTGTTCAAGATTTATTTGGTG 735
 Qy 725 AAAAAAATTTCTGATCAGACCAATTTCTCAGCAACTGTTTATTTACCTTTGGCC 784
 Db 736 AAAAAATTTCTACCAACAATTTCTTTGATCAATTTCTGCCACTCAAGTGTGCTCC 795
 Qy 785 AGTGATTTCTGATCAGATTTTGTAGTAATATCATGTTACTCTGGGTGATTCACACCA 844
 Db 796 GTGAGACTGTAATGAAATTTGAGCAATGCTTTATTTATCATTTTGTGCTTTGACAGG 855
 Qy 845 ACATATGACATGAGCCGAGCAAGTGTATGCTGCCACACTCTTGTGTAACATCTG 904
 Db 856 CAAACTGAAACATGAGTCTGCTGGATGTGTAGTGTGATGATCAATATCCGCGAGGAATTCAG 915
 Qy 905 TGCAAAATTTCTACACTGAGCCAGCAGTGAATTTCTGGTGAATCTCGGGCATTTGACT 964
 Db 916 TTCAAAACATGCTGCTGAGCCAGCTGTTAAATCTGGAATTTTCAAGCTTTTAAT 975
 Qy 965 GGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCCAACTCTCTGTAAGGTACAG 1024
 Db 976 GGGGAAAGTCAGCTCAGAAATGATGCACTTCAATCAGCCCACTCTCCCTACTACAATG 1035
 Qy 1025 TCAGAGATATGAGGCTCCCTACAGCAATGTGACAGAGGTGAGGCTGCTTCCAAATC 1084
 Db 1036 TGACCCCATGATGTGCCAATTTGCAAGTGTGAGTGTGGCAATGACTGTTGGCTGACC 1095
 Qy 1085 CAGAAGACGTGAAATGCTGCTCTGAGGTGACCAACCTCATCTACCAATAGATATTC 1144
 Db 1096 CCAAGATGTGACCTTTGCTTCCAAACTTCTTAATCTCATTTACCAAGGAGATTC 1155
 Qy 1145 CTGAATGGGCTCAGCTGGAATTTCTGCTGGGTTTGGATGCTCTCAGCTATGTAATG 1204
 Db 1156 TTCCATATATCACTTGGATTTTATCTGGCAATGAATGCTCTCAAGAGTTTACAATG 1215
 Qy 1205 AAATCATCCATCTGATGACAGGAGGAGA 1234
 Db 1216 AAATTTTCTATGATGCAAGAGATAA 1245

RESULT 14
 A26690/c
 LOCUS
 DEFINITION Precursor of rabbit gastric lipase (comp.).
 ACCESSION A26690
 1378 bp DNA linear PAT 05-APR-1995

A26690.1 GI:905030
 KEYWORDS synthetic construct
 SOURCE synthetic construct
 ORGANISM artificial sequences.
 REFERENCE 1 (bases 1 to 1378)
 AUTHORS Benicourt, C., Blanchard, C. and Junien, J.L.
 TITLE Recombinant gastric lipase from rabbit and pharmaceutical compositions
 JOURNAL Patent: EP 0542629-A 10 19-MAY-1993;
 INSTITUT DE RECHERCHE JOUVEINAL
 FEATURES Location/Qualifiers
 1..1378
 /organism="synthetic construct"
 /mol_type="genomic DNA"
 /db_xref="taxon:32630"
 BASE COUNT 418 a 276 c 287 g 397 t
 ORIGIN
 Query Match 38.6%; Score 489.2; DB 6; Length 1378;
 Best Local Similarity 65.0%; Pred. No. 8.6e-140;
 Matches 722; Conservative 0; Mismatches 388; Indels 0; Gaps 0;
 Qy 125 ACCCAGAGCATTCATGAATATTAGTGAATATCAATCAATCAAGGCTATCCCTGTGAGG 184
 Db 1243 ACCCTGAAGTGAATATGAATATTAGTCAGATGATTTCTACTGGGATACCCAAAGTGAA 1184
 Qy 185 AATATGAAGTGGCAACTGAAGATGGGTATATCTTTCTGTAAACAGGATTCCTCGAGGCC 244
 Db 1183 AATATGAAGTGTGACTGAAGATGGTATATCTCTGAGGTTAAACAGATTCCTTATGGGA 1124
 Qy 245 TAGTGCAACCTAAGACAGAGGTTCCAGGCTGTGCTGTGTTACTGACGATGGCTAGTTG 304
 Db 1123 AGAAAAATTCAGGAAACAGAGGCCAGAGACCGCTGTATTTTTCAGCAGTGGTTGCTTG 1064
 Qy 305 GAGTGCTAGCAACTGGAATTTCCAACTGCTCCAACTGCTCCAACTGCTGCTTTATTTCTGGCAG 364
 Db 1063 CATCAGCCTCAAACTGATCTCCAACTGCTCCAACTGCTGCTGCTTTATTTCTGGCAG 1004
 Qy 365 ATGCTGTTTTCAGCTGTGATGGGAAACAGAGGAGGAAACGCTGCTCGAAACACA 424
 Db 1003 ATGCTGTTTTCAGCTGTGATGGGAAACAGAGGAGGAAATACCTGCTCGAAAGAACT 944
 Qy 425 AGACACTCTCATAGACCAAGATGAGTTCTGGCTTTTCACTTTATGATGAGATGGCTAGGT 484
 Db 943 TATACTATTACCAAGCTCCGTTGAATTTCTGGCTTTTATGATGAAATGGCTAAT 884
 Qy 485 TTGACCTTCTGAGTGAATAAATTTATTTTTCAGAAACAGGAGGAGGAAAGATCTATT 544
 Db 883 ATGACCTTCCAGCCCAAAATTTGACTTTCATTTGAAAGGAACTGGACAGGAAAGCTGCACT 824
 Qy 545 ATGTCGGCTATTACAGGGGACCACTGCTTTTATTTGCTTTTCCATTTTCCATCCAGAGC 604
 Db 823 ATGTTGTCATTTCCAGGGGACCACTGCTTTTATTTGCTTTTATTTGCTTTTCTACAAATCCCAAGC 764
 Qy 605 TGGCTCAGAAATCAAAATGATTTTTCCTTTAGCACCCTAGCCCTGTTAAGCATGCAA 664
 Db 763 TGGCTGAAGAAATCAAAACCTTTTATGCTAGCTCCAGTCGCACTGTGAAGTATACAA 704
 Qy 665 AAAGCCCGGAGCAAAATTTTCTGCTGCAATGATGATGATGATGATGATGATGATGATGATG 724
 Db 703 AAAGCCCTGTAACAAACTTAGGTTTTATTTCTCAACATGTTTCAAGATTTATTTGGTG 644
 Qy 725 AAAAAGAAATTTCTGATCAGACCAAGATTTCTCAGACCAACTTTGTTATTTTACCTTTTGGCC 784
 Db 643 ACAAATATTCTACCCACACAATTTCTTTGATCAATTTCTTTGCTGCTCAAGTGTGTTCC 584
 Qy 785 AGGTGATCTTGTAGTGAATTTGTAGTAATATGATGATGATGATGATGATGATGATGATGATG 844
 Db 583 GTGAGACACTGAATGTAATTTGAGCAATGCTTTATTTATTTATTTGCTGCTTTGACAGG 524
 Qy 845 ACAATATGAACATGAGCCGAGCAAGTGTATGCTGCCACACTTTGCTGGAACATCTG 904

```

sig_peptide
45..101
/ISOECS="I"
/note="pot. signal peptide (AA -19 to -1)"
CDS
63..1241
/codon_start=1
/product="gasic lipase precursor"
/protein_id="CAA29414.1"
/db_xref="GI:758064"

```

```

/db_xref="SWISS-PROT:P07098"
/translation="MASISVLGTHGLFGKLPKHSPEVTVNLSQMTIYGYNPEYNE
VTEDEYILLENRIPIYKKNSGNTQRPVFLQVGLLASATNINSLPNNSLAFILAL
HYVSGWNGTTCGFIATFSPSLAKRIKTYSPDSVEFWAFSEDEMAKTDLPATIDFIYKTKQOKL
AGYHDSQSGTTCGFIATFSPSLAKRIKTYSPDSVEFWAFSEDEMAKTDLPATIDFIYKTKQOKL
IFGDKTIFYNHPFDQFLATKVCSEMLNLCNALFIICGFDSDKNFTNSTRLDVYLSHN
PAGTSVQNFHEFTQAVKSGKQAPQSCVQNCNRYHVDQSPPYVNTAMNVPJAVNGN
GKOLLADPDQDVGILLPKLPNLIIYKELIPFYNHLDFIWMADAPQEVYNDIVSMISEDKK
"
sig_peptide      63..101
                  /notes="pot. signal peptide (AA -13 to -1)"
mat_peptide      102..1238
                  /product="gastric lipase (AA 1-379)"
BASE_COUNT      396 a 284 c 268 g 417 t
ORIGIN
Query Match      37.0%; Score 470; DB 9; Length 1365;
Best Local Similarity 64.2%; Pred. No. 7.8e-134;
Matches 707; Conservative 0; Mismatches 395; Indels 0; Gaps 0;
QY 119 CTGTGGACCCAGAGCAATTCATGAATATATAGTGAATCATCCACATCAAGGCTATCCCT 178
Db 121 CTGGAAGCCCTGAAGTGAATATGAACATTTAGTCAGATGATTACTTATTGGGGGATACCCAA 180
QY 179 GTGAGGAATATGAAGTCGCAACTGAAGATGGGTATATCCTCTTGTGTTAAACAGGATTCCT 238
Db 181 ATGAAGAATATGAAGTTGTGACTGAAGATGGTTATATCTTTGAAGTCAATAGATTCCTT 240
QY 239 GAGGCGTAGTGCAACCTTAAGAGACAGGTTTCAGGCGCTGTGGTGTACTGCAGCATGGCC 298
Db 241 ATGGGAAGAAATAATTCAAGGATATACAGGCGCAGACCTGTGTGTGTTTTGCGCATGGTT 300
QY 299 TAGTTGCGAGTGTGTAGCACTGGATTTCCAACTGTCGCCAAATAATAGCCTGGGCTTCATTC 358
Db 301 TCGTTGCATCAGCCACAACTGGATTTCCAACTGCGGNAACAAAGCCCTTGCCTTCATTC 360
QY 359 TGGCAGATGCTGGTTTGTGAGTGTGATGGGAAACAGCAGGGGAAACCGCTTGCTCGAA 418
Db 361 TGGCAGATGCTGGTTATGATGTGTGCTGGGCAACAGCAGAGGAAACACCTGGGCGCAGAA 420
QY 419 AACCAAGACACTCTCCATAGCAAGATGAGTCTGTGGCTTTTCAGTTTATGATGAGATGG 478
Db 421 GAAACTTGTAATTATCACCAATTCAGTTGAAATTTCTGGGCTTTTCAGCTTTGATGAAATGG 480
QY 479 CTAGGTTTGACCTTCTCTGCAGTGATAAACTTTATTTTGCAGAAAAACGGGCCAGGAAAGA 538
Db 481 CTAATATGACCTTCCAGCCACAACTCGACTTCATTGTAAAGAAACCTGACAGAAAGCAGC 540
QY 539 TCTATATATGTCGGCTATTTCAGGGGACCAACCATGCGGCTTTATTCGATTTTCCACCATGC 598
Db 541 TACACTATGTGTGGCCATTCCTCCAGGGCACCAACATGTTTATATGCTCTTTTCCACCAATC 600
QY 599 CAGAGCTGGCTTCAGAAAAACAAATATGTTATTTGCTTTAGCACCCATAGCCACTGTTAAGC 658
Db 601 CCAGCTGGCTTAAAGAAATCAAAACCTTCTATGCTCTAGCTCTCTGTGCCACTGTGAAGT 660
QY 659 ATGCAAAAGCCCGGACCAAAATTTTGTGTGTCGCAATATGATGATCAAGGATGTT 718
Db 661 ATCAAAAGCCCTTATAAACAAACTTAGATTTGTTTCTCAATCCCTCTTCAAGTTTATAT 720
QY 719 TTGCAAAAAGAGATTTCTGTATCAGACAGATTTCTCAGACAACTGTGTTATTTACCTTT 778
Db 721 TTGTGTGCAAAATATTTCTACCCACAACTTCTTTGATCAATTTCTTGCTACTGAAGTGT 780
QY 779 GTGGCCAGGTGATTTCTTGATCAGATTTGTAGTAATATCATGTTACTTCTGGGTGATTC 838
Db 781 GCTCCCGTGATGCTGAAATCTCCTTTGAGCAATGCCCTTATTTATATTTTGTGATTTG 840
QY 839 ACACCAACATATGAAACATGACCGGACGAGTGATATGCTGCCACACTCTTGCTGCGAA 898
Db 841 ACAGTAAGAACTTTAAACACAGATGCTGTGATGTATCTATCAATATCAATCAGCAGAA 900
QY 899 CATCTGTGCAAAATATTTCTACATCTGAGCGCAGGACGTGAATTTCTGGTGAATCCGGGAT 958

```

Db	901	CTTCTGTTCAAAACATGTTCCATTGGACCCAGGCTGTTAAGTCTGGGAAATTCACAGCTT	960
Qy	959	TTGACTGGGGAGTGAGACCAAAATCTGGAAAAATGCAATCAGCCAACTCCCTGTAAGGT	1018
Db	961	ATGACTGGGAAGCCCAAGTTCAAGATAGGATGCACTATGATCAGTCCCAACCTCCCTACT	1020
Qy	1019	ACAGAGTCAGAGATATGACGGTCCCTACAGCAATGTGGACAGGAGGTGAGGACTGGGCTTT	1078
Db	1021	ACAATGTGACAGCCCATGATGTACCAATTGCAATGTGGACGGTGGCAAGGACCTGTTGG	1080
Qy	1079	CAATCCAGAGAGCGTGAATGCTGCTCTGTGAGGTGACCACTCATCTACCATTAAGA	1138
Db	1081	CTGACCCCAAGATGTTGGCCTTTTGCTTCCAAAACCTCCCAATCTTATTATACCAAGG	1140
Qy	1139	ATATTCCTGAATGGGCTCACGTGGATTTTCATCTGGGGTTTGGATGCTCCTCACCGTATGT	1198
Db	1141	AGATTCTCTTTTACAAATCACTTGGACTTTATCTGGCAATGATGCCCTCAAGAAGTTT	1200
Qy	1199	ACAATGAATCATCCATCTGAT	1220
Db	1201	ACAATGACATTGTTTCTATGAT	1222

Search completed: February 19, 2004, 08:29:21
Job time : 4672.38 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 04:48:12 ; Search time 2775.48 Seconds
(without alignments)
11112.438 Million cell updates/sec

Title: US-10-042-431-46

Perfect score: 1269

Sequence: 1 atgttggaacacctgtcaag.....gacgggtgtgagccgattg 1269

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: em_estba.*

2: em_esthum.*

3: em_estin.*

4: em_estmu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_htc.*

9: gb_estl.*

10: gb_est2.*

11: gb_htc.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_eston.*

17: em_gss_hum.*

18: em_gss_inv.*

19: em_gss_pln.*

20: em_gss_vrt.*

21: em_gss_fun.*

22: em_gss_mam.*

23: em_gss_mus.*

24: em_gss_pro.*

25: em_gss_rod.*

26: em_gss_phg.*

27: em_gss_vrl.*

28: gb_gssl.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	961.8	75.8	2050	11 AK037091	AK037091 Mus muscu
2	961.8	75.8	2678	11 AK085719	AK085719 Mus muscu
3	960.2	75.7	2555	11 AK037214	AK037214 Mus muscu
4	618.2	48.7	2927	11 AK019504	AK019504 Mus muscu

5	505	39.8	954	10	BF135102	BF135102 601780142
6	497.6	39.2	1079	10	BF141567	BF141567 601788654
7	475.4	37.5	2138	11	AK088659	AK088659 Mus muscu
8	471.8	37.2	689	10	BF143546	BF143546 601790896
9	461.8	36.4	1350	11	AK010093	AK010093 Mus muscu
10	460.2	36.3	1347	11	AK010116	AK010116 Mus muscu
11	460.2	36.3	1349	11	AK009300	AK009300 Mus muscu
12	460.2	36.3	1350	11	AK010139	AK010139 Mus muscu
13	458.6	36.1	1340	11	AK010124	AK010124 Mus muscu
14	458.6	36.1	1343	11	AK009479	AK009479 Mus muscu
15	458.6	36.1	1345	11	AK009413	AK009413 Mus muscu
16	458.6	36.1	1345	11	AK009473	AK009473 Mus muscu
17	458.6	36.1	1346	11	AK009359	AK009359 Mus muscu
18	458.6	36.1	1346	11	AK009523	AK009523 Mus muscu
19	458.6	36.1	1346	11	AK009729	AK009729 Mus muscu
20	458.6	36.1	1348	11	AK010058	AK010058 Mus muscu
21	458.6	36.1	1349	11	AK009474	AK009474 Mus muscu
22	458.6	36.1	1349	11	AK009546	AK009546 Mus muscu
23	458.6	36.1	1349	11	AK009571	AK009571 Mus muscu
24	458.6	36.1	1349	11	AK009573	AK009573 Mus muscu
25	458.6	36.1	1349	11	AK010148	AK010148 Mus muscu
26	458.6	36.1	1350	11	AK009459	AK009459 Mus muscu
27	458.6	36.1	1350	11	AK009525	AK009525 Mus muscu
28	458.6	36.1	1350	11	AK009773	AK009773 Mus muscu
29	458.6	36.1	1350	11	AK010061	AK010061 Mus muscu
30	458.6	36.1	1350	11	AK010125	AK010125 Mus muscu
31	458.6	36.1	1350	11	AK010203	AK010203 Mus muscu
32	458.6	36.1	1351	11	AK009428	AK009428 Mus muscu
33	458.6	36.1	1352	11	AK009431	AK009431 Mus muscu
34	458.6	36.1	1356	11	AK010019	AK010019 Mus muscu
35	458.6	36.1	1356	11	AK010035	AK010035 Mus muscu
36	457.6	36.1	1346	11	AK009437	AK009437 Mus muscu
37	457.6	36.1	1347	11	AK010103	AK010103 Mus muscu
38	457.6	36.1	1349	11	AK010026	AK010026 Mus muscu
39	457	36.0	1346	11	AK009560	AK009560 Mus muscu
40	457	36.0	1347	11	AK010236	AK010236 Mus muscu
41	457	36.0	1348	11	AK010231	AK010231 Mus muscu
42	457	36.0	1349	11	AK009875	AK009875 Mus muscu
43	457	36.0	1349	11	AK009537	AK009537 Mus muscu
44	456	35.9	1349	11	AK010106	AK010106 Mus muscu
45	455.4	35.9	1345	11	AK009544	AK009544 Mus muscu

ALIGNMENTS

RESULT 1

AK037091

LOCUS

DEFINITION

AK037091

2050 bp

linear

mus musculus adult female vagina cDNA, RIKEN full-length enriched library, clone:9930111004 product:BA30415.1 (NOVEL LIPASE)

(FRAGMENT) homolog [Homo sapiens], full insert sequence.

ACCESSION

AK037091

VERSION

AK037091.1

KEYWORDS

HTC; CAP trapper.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE

1

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

MEDLINE

99279253

PUBMED

10349636

REFERENCE

2

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE

20499374

PUBMED

11042159

REFERENCE
AUTHORS

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

TITLE

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamahara, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G.,
Quackenbush, J., Schriml, L.M., Stauber, R., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boiffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
Toyooka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kotsuki, S.
and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217951

TITLE

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

5 The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2050)

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, T., Tagawa, A., Takahashi, P., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission

TITLE

JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers

FEATURES

source 1. .2050
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:9930111004"
/db_xref="taxon:10090"
/clone="9930111004"
/sex="female"
/tissue_type="vagina"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
144. 1412
CDS
/note="unnamed protein product; BA30415.1 (NOVEL LIPASE)
(FRAGMENT) homolog (Homo sapiens) [SPTR|Q96LG2, evidence:
FASTY, 88.5%ID, 100%length, match=1119]
putative"
/codon_start=1
/protein_id="BAC29699.1"
/db_xref="GI:26331938"
/translation="MSEIISRVVTVSHRVEIMLLIIVLYLQNVNNSHLPKAAADPE
AFNVSEI IKHGPSPSEYEVEDGYILSVNRIPRGQTLKKEGRPVVLLQHLGLG
DASWISLNPNSLIGFILADAGDVGNGSRGNTSRKHTLSIDQEFWAFSYDMA
RFDLPVINFILQKTKQKVVYVYSGQTMGFIAFTWPELAHKIMYFALAPIATV
KYAESPTKELLIPDMKIVLFGRELCTPFRCLFYLQGMILDOICSNIIILL
GGNTNNNSRANVVAHVPAGTSVQNLHWSQVNSBELRAFDWGSSETKQEKCNQ
PTPIRYKVRDMVPTAMWTGGQWLSNPDVKTLLSEVNLIIHKNIPEWAHVDFIHW
LDAPQRYNEIHLHMQEPLPQGTCTRVKL"
2027. .2032
polya_signal /note="putative"
2050
polya_site /note="putative"
BASE COUNT 598 a 415 c 461 g 576 t
ORIGIN
Query Match 75.8%; Score 961.8; DB 11; Length 2050;
Best Local Similarity 85.4%; Pred. No. 1.5e-279;
Matches 1084; Conservative 0; Mismatches 182; Indels 3; Gaps 1;
QY 1 ATGTGGAAACCTTGTCAAGACAGTGGATGTCTCACAGAGATGGAATGTGGCTTCG 60
DB 144 ATGTGAGAAATCTTGTCAAGAGTGTGGACTGTTTCGCACAGAGTGGAGATGGCTTCG 203
QY 61 ATTCGTGGGGTATATGTTCCAGAGAAATGTGAATTCAGTACATATGCCAACTAAAGCT 120
DB 204 ATTCGTGAGTATTTACTTCCAGAAATGTGAATTCGGACATTTGCCACGAAAGCT 263
QY 121 GTGACCCAGAAAGCATTCATGAATATTAGTGAATATCCAAATCAAGCTATTCCTGT 180
DB 264 GCGGATCCAGAAAGCATTCATGAATATTAGGCAAAATCATCAACACAAAGGGTTATCCAGT 323
QY 181 GAGGAATATGAAGTCGCAACTGAAGATGGGTATATCTTCTGTAAACAGATTCTTCGA 240
DB 324 GAGGAGTATGAAGTTGCAACGAAAGTGGTATCATCTTCTGTGAACAGAAATCCCTCGG 393
QY 241 GGCCTAGTGCACCTTAAGAGACAGGTTCACAGGCTGTGGTGTATCTGACGATGGCCTA 300
DB 384 GGACAGACACGGTTAAAGAGAGAGAGTCCAGGCCAGTGGTGTACTGCAACATGGTCTT 443
QY 301 GTTGAGGTGTAGCAACTGGATTTCCAACTGCCCAACAATAGCTGGGCTTCATTCG 360
DB 444 CTTGGGATGTAGCAACTGGATTTCCAACTGCCCAACAACAGCTGGGCTTTATTCG 503
QY 361 GCAGATGCTGGTTTTCAGCTGTGGATGGGAAACAGCAGGGGAAACGCTGCTTCGAAAA 420
DB 504 GCAGATGAGGTTTGTGATGTGTGGATGGGAAACAGCAGAGGAAACACTGGTCTCGAAG 563
QY 421 CACAAGACATCTCCATAGACAAGATGAGTTCTGGGCTTTCAGTTATGATGAGATGGCT 480
DB 564 CACAAGACCTCTCCATAGATCAAGATGAGTTCTGGGCTTTCAGTTATGATGAGATGGCT 623
QY 481 AGTTTGAACCTTCTCGAGTGTATAAATTTATTTTGCAGAAACGGGCCAGGAAAGATC 540
DB 624 AGTTTGAACCTTCCAGCTGTGTATAAATTTATTTATCAGAAACGGGCCAGGAAAGATC 683

QY 541 TATTATCGGCTATTCACAGGCGCACCATGGCTTTATTCGATTTTCCACCATGCCA 600
 Db 684 TATTATCGGCTATTCACAGGCGCACCATGGCTTTATTCGATTTTCCACCATGCCA 743
 QY 601 GAGCTGGCTCAGAAATCAAAATGATTTTGGCTTTAGCAGCCCATAGCACTGTTAAGCAT 660
 Db 744 GAGCTAGCTCATAAATCAAAATGATTTTGGCTTTAGCAGCCCATAGCACTGTTAATAT 803
 QY 661 GCATAAAGCCCGGACCAATTTTGTCTGCGCAGATGATGATCAAGGATTTGTT 720
 Db 804 GCAGAGAGTCTGTGATCAAAATTTCTGCTGCGCAGATGATGATCAAGGATTTATTT 863
 QY 721 GGCATAAAGAAATTTCTGATCAGACAGATTTCTCAGACAACTTGTATTATACCTTTGT 780
 Db 864 GGCAGACAGAGTTTATACAGACAGATTTTCTCAGACAGCTTTTATTTACCTTTGT 923
 QY 781 GGCAGAGTATCTGATCAGATTTGATGATTAATCATGTTCTGCTGGTGAATCAAC 840
 Db 924 GGCAGAGTATCTGATCAGATTTGATGATTAATCATGTTCTGCTGGTGAATCAAC 983
 QY 841 ACCAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
 Db 984 ACCAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1043
 QY 901 TCTGTGCAAAATATTTCTGATCAGACAGATTTCTGATGATGATGATGATGATGATGATGAT 960
 Db 1044 TCTGTGCAAAATATTTCTGATCAGACAGATTTCTGATGATGATGATGATGATGATGATGAT 1103
 QY 961 GACTGGGGAGTGAACCAAAATCTGAAATGATGATGATGATGATGATGATGATGATGATGAT 1020
 Db 1104 GACTGGGGAGTGAACCAAAATCTGAAATGATGATGATGATGATGATGATGATGATGATGAT 1163
 QY 1021 AGACTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
 Db 1164 AAGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1223
 QY 1081 AATCAGAGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
 Db 1224 AATCAGAGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1283
 QY 1141 ATTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
 Db 1284 ATTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1343
 QY 1201 AATGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
 Db 1344 AATGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1400
 QY 1261 GCGGTATTG 1269
 Db 1401 GTCAAATG 1409

RESULT 2
 LOCUS AK085719
 DEFINITION Mus musculus 10 days lactation, adult female mammary gland cDNA,
 RIKEN full-length enriched library, clone:D730026117
 product:BA30415.1 (NOVEL LIPASE) (FRAGMENT) homolog [Homo sapiens],
 full insert sequence.

ACCESSION AK085719
 VERSION AK085719.1 GI:26351760
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1
 Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253

10349636

PUBMED
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20499374
 11042159

3
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

JOURNAL
 MEDLINE
 PUBMED

20530913
 11076861

4
 REFERENCE
 AUTHORS

TITLE

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.go.jp/>

URL: <http://fantom.gsc.riken.go.jp/>

FEATURES

source

```
1..2678
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:D730026117"
/db_xref="taxon:10090"
/clone="D730026117"
/sex="female"
/tissue type="mammary gland"
/clone lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 days lactation, adult"
118..1386
/notes="unmated protein product; BA30415.1 (NOVEL LIPASE) (FRAGMENT) homolog [Homo sapiens] (SPR|096Lg2, evidence: FASTV, 88.5%id, 100%length, match=1119) putative"
```

CDS

```
/codon_start=1
/protein_id="BAC39517.1"
/db_xref="GI:26351761"
/translation="MSEILSRVWTVSHRVEIWLILVAILQNVNSGHLPTKAADPE
AFNVSEIHKHGYDSEEVATEDGYILSVNRIERGQRLKKGSRVVLQGLLG
DASNIINLPNLSGFIADAGFDVWNGNSRNTWSRHKHLSIDQDSFWAFSDEMA
RDLPAVINFLQKTQKKVYVYVQSQTWGFIAFSPPELAHKIKMYFALAPIATV
KVARSPTKFLPLPMKIVLPGREFLYQTFRFLQFLYLCGMILQDQISNILLIL
GGFTNNMMSRANVVAHTPAGTSVQNLHWSOAVNSGELRAPDWSGETKNQKCNQ
PPPIKVPZDMVPTAMWVGQDLSNPDDVKTLLSEVNTLYHKNIPFWAHVDFING
LQAPQVNEIHLMKQENLPQGCIVKL"
```

polyA_signal

```
2657..2662
/notes="putative"
```

polyA_site

```
2678
/notes="putative"
```

BASE COUNT 795 a 522 c 555 g 806 t

ORIGIN

```
Query Match 75.8%; Score 961.8; DB 11; Length 2678;
Best Local Similarity 85.4%; Pred. No. 1.8e-279;
Matches 1084; Conservative 0; Mismatches 182; Indels 3; Gaps 1;

Qy 1 ATGTTGAAACCTTGTCAAGACAGTGGATTGTCTCACAGAGATGGAATGGGTTCTG 60
Db 118 ATGTCAGAAATCTTGTCAGAGTGTGGACTGTTCGTCAGAGTGGAGATGGGTTCTG 177

Qy 61 ATTCTGGTGGCGTATATGTTCCAGAGAAATGTGAATTCAGTACATATGCCAACTAAAGCT 120
Db 178 ATTCTGGTAGCGTATTTACTCCAAAGAAATGTGAATTCGGGACATTTGCCACGAAAGCT 237

Qy 121 GTGAGCCAGACAGCTCATATATAGTGAATCATCCACATCAAGGCTATCCCTCT 180
Db 238 GCGGATCCAGAGCATTCATGAATGTAGGGAATCATCAACAGAGGTTATCCAGT 297

Qy 181 GAGGAATATGAAGTCGCAACTGAAGATGGGTATATCTTCTGTGTTAAGAGATTCCTCGA 240
Db 298 GAGGAGTATGAAGTTGCAACGGAAGATGGGTATATCTTCTGTGTTAAGAGATTCCTCGG 357

Qy 241 GGCTAGTGCACCTAAGAGACAGGTTCCAGGCTGTGGTGTACTGCAGCATGCCTTA 300
Db 358 GGACAGACACGGTTAAAGAGAGGAGGATCCAGGCCAGTGGTGTACTGCACATGGTCTT 417

Qy 301 GTTGGAGGTGTAGCACTGGATTTCACACTGCCCAACAAATAGCTGGGCTTCATCTG 360
Db 418 CTTGGGATGCTAGCACTGGATTTCACACTGCCCAACAAATAGCTGGGCTTCATCTG 477

Qy 361 GCAGATCTGGTTTGTAGCTGTGGATGGGGAACAGAGGGGAAACGCTGGTCTCGAAA 420
```

```
Db 478 GCAGATGCAGGTTTTCATGTGTGGATGGAAACACAGCAGGAAACACCTGGTCTCGAAG 537
Qy 421 CACAAGACACTCTCCATAGACCAAGATGAGTCTTGGGCTTTCAGTATGATGAGATGGCT 480
Db 538 CACAAGACCTCTCCATAGATCAAGATGAGTCTTGGGCTTTCAGTATGATGAGATGGCT 597
Qy 481 AGGTTTGAACCTCTCTGACGTGATAAACTTTATTTTGCAGAAAACGGGCGAGAAAGATC 540
Db 598 AGGTTTGAACCTCTGACGTGATAAACTTTATTTTGCAGAAAACGGGCGAGAAAGATC 657
Qy 541 TATTATGTCGGCTATTCACAGGGCCACCATGGGCTTTTATGCAATTTTCCACCATGCCA 600
Db 658 TATTATGTCGGCTATTCACAGGGCCACCATGGGCTTTTATGCAATTTTCCACCATGCCA 717
Qy 601 GAGCTGGCTCAGAAAATCAAAATGATTTTGTAGCACCACCATAGCCACTGTGAAGCAT 660
Db 718 GAGCTAGCTCATTAATCAAAATGATTTTGTAGCTTCTATAGCCACTGTGAATAT 777
Qy 661 GCAAAAGCCCGGGACCAAAATTTTGTGTCGCGAGATGATGATGATCAAGGATGTTT 720
Db 778 GCAAAAGCTCTGGTACCAAAATTTCTGCTGCTGCCAGATGATGATGATCAAGGATGTTT 837
Qy 721 GGCAGAAAAGAAATTTCTGTATCAGACCATTTCTCAGACCAACTTGTATTTTACCTTCT 780
Db 838 GGCAGAGCAAGAGTTTTTATACCAAGATGATTTTTCAGACAGCTTTTATTTTACCTTCT 897
Qy 781 GGCAGAGTATTTCTGATCAGATTTTGTAGTAATATCATTTTCTGCGGTGGATTCAC 840
Db 898 GGCAGAGTATTTCTGACCAAAATCTCAGCAACATCATCTTACTCTCGGAGGTTTTTAC 957
Qy 841 ACCAACAATATGAACTGAGCCGAGCAAGTGTATATGCTGCCACACTCTTGTCTGGAACA 900
Db 958 ACAACAATATGAACTGAGCAGAGCAAAATGTGTGCTGCCCATACACTGCTGGAACG 1017
Qy 901 TCTGTGCAAAATATTTCTACACTGGAGCCAGGAGTGAATTTCTGGTGAATCTCCGGGCA 960
Db 1018 TCTGTGCAAGAACATTTCTCACTGGAGCCAGGCGGTGAATTTCTGGGGAATCTTGTG 1077
Qy 961 GACTGGGGAGTGAGACCAAAATCTGGAATAATGCAANTCAGCCAACTCTGTAAAGTAC 1020
Db 1078 GACTGGGGAAGTGAGACCAAAATCAGGAGAAATGTAANTCAGCCAACTCTTATAAGGTAC 1137
Qy 1021 AGAGTCAGAGATATGACGGTCCCTCAGCAATGTGACAGGAGGTGAGGACTGGCTTCA 1080
Db 1138 AAGTTTCAGATATGATGTGCTCCAAACAGCAATGTGAGTGGAGTCAAGACTGGCTTCA 1197
Qy 1081 AATCCAGACAGCTGGAATAATGCTGCTCTGAGGTGACCAACCTCATCTACCAAGAAG 1140
Db 1198 AATCCAGATGATGGAATAATGCTGCTCTGAGGTGACCAACCTCATCTACCAAGAAG 1257
Qy 1141 ATTCTGAAATGGGCTCAGTGGATTTTCATCTGGGTTTGGATGCTCTCACCGTATGTAC 1200
Db 1258 ATTCTGAGTGGGCTCATGTGGATTTTCATCTGGGCTGGATGCCCTCAGCGTGTATAC 1317
Qy 1201 AATGAAATCATCCATCTGATGTCAGAGGAGGAGACCAACTTTTCCAGGAGCGGTGTAG 1260
Db 1318 AATGAAATTTATCATCTGATGATGAAGC---AGGAGCCCAACCTTCCCGAGGAACCTGCAGG 1374
Qy 1261 GCCGTATTG 1269
Db 1375 GTCAAAATTG 1383
```

RESULT 3

AK037214

LOCUS

DEFINITION

AK037214

ACCESSION

VERSION

KEYWORDS

SOURCE

AK037214 2555 bp mRNA linear HTC 05-DEC-2002
Mus musculus 6 days neonate skin cDNA, RIKEN full-length enriched
library, clone:AK030010N08 product:BA30415.1 (NOVEL LIPASE)
(FRAGMENT) homolog [Homo sapiens], full insert sequence.

AK037214.1 GI:26332053

HTC; CAP trapper.

Mus musculus (house mouse)

ORGANISM	Mus musculus	TITLE	Direct Submission
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	JOURNAL	Submitted (15-JUL-2001) Yoshihide Havaehizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
AUTHORS	Carninci, P. and Hayashizaki, Y.	COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
TITLE	High-efficiency full-length cDNA cloning	URL: http://genome.gsc.riken.go.jp/	
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)	URL: http://fantom.gsc.riken.go.jp/	
MEDLINE	99279253	Location/Qualifiers	
PUBMED	10349636	1. 2555	
REFERENCE		/organism="Mus musculus"	
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.	/mol_type="mRNA"	
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	/strains="C57BL/6J"	
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)	/db_xref="FANTOM_DB:A030010N08"	
MEDLINE	20499374	/db_xref="taxon:10090"	
PUBMED	11042159	/clone="A030010N08"	
REFERENCE		/tissue_type="skin"	
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagasaka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.	/clone_lib="RIKEN full-length enriched mouse cDNA library"	
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multichannel sequencer	126. 1394	
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)	/note="unnamed protein product; BA30415.1 (NOVEL LIPASE) (FRAGMENT) homolog [Homo sapiens] (SPTR Q96LG2, evidence: FASTA, 88.5%ID, 100%length, match=1119)	
MEDLINE	20530913	putative"	
PUBMED	11076861	/codon_start=1	
REFERENCE		/protein_id="BAC29757.1"	
AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Asbunier, M., Batalov, S., Casavant, T., Fleischman, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Kachiwa, H., Quackenbush, J., Schrim, L.M., Staubli, F., Suzuki, G., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Anono, H., Baldaire, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Garibolli, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kaniwa, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombarts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyo-oka, K., Wang, K.H., Weitz, C., Wittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.	/dev_stage="6 days neonate"	
TITLE	Functional annotation of a full-length mouse cDNA collection	BASE COUNT	752 a 501 c 536 g 766 t
JOURNAL	Nature 409 (6821), 685-690 (2001)	ORIGIN	
MEDLINE	21085660	Query Match	75.7%; Score 960.2; DB 11; Length 2555;
PUBMED	11217851	Best Local Similarity	85.3%; Pred. No. 5.4e-279;
REFERENCE		Matches 1083; Conservative	0; Mismatches 183; Indels 3; Gaps 1;
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.	QY	1 ATGTTGGAAACCTTGTCAAGACAGTGGATTCTCTCACACAGATGGAATCGCTTCTG 60
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	Db	126 ATGTGAGAAATCTTGTCAAGAGTGTGGACTGTTTCGCACAGTGGAGATATGCTTCTG 185
JOURNAL	Nature 420, 563-573 (2002)	QY	61 ATTCTGGTGGCGTATATGTTCCAGAGAAATGGAATTCAGTACATATATCCCAACTAAAGCT 120
REFERENCE		Db	186 ATTCTGGTAGCGTATTTACTCCAAAGAAATGTGAATCGGACATTTGCCCAACCAAGCT 245
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, P., Imokani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, F., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.	QY	121 GTGGACCCAGAGCAATTCATGAATATAGTGAATATCCCAATCAAGCTATCCCTGT 180
TITLE		Db	246 GCGGATCCAGAGCAATTCATGAATGTAGCGAAATCATCAACACCAAGGGTTATCCCACT 305
JOURNAL		QY	181 GAGGATATGAAGTCCGAACCTGAAGATGGGTATATCTCTTTCTGTAAACAGGATTCCTCGA 240
REFERENCE		Db	306 GAGGATATGAAGTGTGCAACCGAGATGGGTACTCTCTTTCTGTGAACAGATCCCTCGG 365
AUTHORS		QY	241 GGCCTAGTGCAACCTAAGAGCAGGTTCAGAGCTCTGGTGTGTACTGACAGCATGGCCTA 300
TITLE		Db	366 GGACACACACCGTGTAAAGAGAGAGGATCCAGGCCAGGTGGTGTACTGCAACATGGTCTT 425
JOURNAL		QY	301 GTTGGAGGTGTAGCAACTGGATTTCCAACTGCCCAACAATAGCTGGGGCTTCAATCTG 360
REFERENCE		Db	426 CTTGGGATGTAGCAACTGGATTTCCAACTGCCCAACAACAGCTGGGGTCTTATCTG 485


```

/clone_lib="NCI_CGAP_Lu30"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; transgenic model WNT-1, expression driven by
MTMV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
dT. Library constructed by Life Technologies.
Investigator providing samples: Gilbert Smith, NIH"
BASE COUNT      280 a  222 g  222 c  222 g  230 t
ORIGIN

Query Match      39.8%; Score 505; DB 10; Length 954;
Best Local Similarity 85.3%; Pred. No. 3.2e-141;
Matches 598; Conservative 0; Mismatches 100; Indels 3; Gaps 3;

QY 493 CTGCGAGTGATAAATTTATTTTGCAGAAAACGGGCCGAGAAAAGATCTATTATGTCGCG 552
DB 1  CGAGCTGTGTAAATTTGATCTTACAGAAAACGGGCCGAGAAAAGTCTATTATGTCGCG 60

QY 553 TATTCACAGGGACACCAATGGGCTTTATGTGATTTTCCATGCGCAGAGCTGGCTGAG 612
DB 61 TACTCACAGGGACACCAATGGGCTTTATGTGATTTTCCATGCGCAGAGCTAGCTCAT 120

QY 613 AATATCAAAATGATTTTGTCTTGTAGCAGCATAGCCACTGTTAAGCATGCAAAAAGCCOC 672
DB 121 AATATCAAAATGATTTTGTCTTGTAGCAGCATAGCCACTGTTAATATGCAAAAGTCT 180

QY 673 GGGACCAAAATTTTGTCTGCGCAGATATGATCAAGGGATTTGTTGGCAAAAAGAA 732
DB 181 GGTACCAAAATTTCTGTCTGCGCAGATATGATCAAGGTATTTATTTGGCAGACAAGAG 240

QY 733 TTTCTGTATCAGACCAAGATTTCTCAGACAACTTTGTTTACTTTTGTGGCCAGGTGATT 792
DB 241 TTTTATACCAAGACTAGATTTTTCAGGAGCTTTTATTTTACC-TTGTGGCCAGATGATT 299

QY 793 CTTCATCAGATTTGTGTATATATCATGTTACTTCTGGGTGGATTTCAACACCAACATATG 852
DB 300 CTCGACCAATCTGCAGCAACATCATCTTACTCTCTGGAGGTTTAAACACAAACATATG 359

QY 853 AACATGAGCCGAGAGATGTATATGTGCCCAACACTTTGCTGGGAACATCTGTGCAAAAT 912
DB 360 AACATGAGCAGAGCAAAATGTATGTGCGCCATACACCTGCTGGAAACGCTGTGCGAAG 419

QY 913 ATTCTACATCTGGAGCAGGAGTGAATTTCTGTGTGAATCTCGGGCATTTGACTGGGGAGT 972
DB 420 ATTCTACATCTGGAGCAGGCGGTGAATTTCTGGGGAACCTTCTGCTTGGGGAAT 479

QY 973 GAGACCAAAATCTGGAAAATGCAATTCAGCCAACTCTCTGAAGGTACAGATTCAGAGAT 1032
DB 480 GAGACCAAAATCAGGAGAAATGTAATTCAGCCAACTCTCTAAGGTACAAAGTTCAGAT 539

QY 1033 ATGACGCTCCCTACAGCAATGTGACAGAGGTCAGGACTGGCTTTCAAAATCCAGAGAC 1092
DB 540 ATGATGTCCCAACAGCAATGTGACCTGGAGGTCAAGACTGG-TTTCAAATCCAGATGAT 598

QY 1093 GTGAAAATGTCTCTCTGAGGTGACCAACTCATCTACCATAGAATATTCCTGAATGG 1152
DB 599 GTGAAAACAAATTACTTTCTGAAGTAACCAAACTCATCTACCAAGAACATTCCTGAGGTG 658

QY 1153 GCTCACGTGATTTTCATCTGGGTTTGGATGCTCTCTCACCG 1193
DB 659 GGTATGTGGA-TTCATCTGGGGGGTGGATGCCCCCTCAGCG 698

RESULT 6
LOCUS      BF141567
DEFINITION 601788654F1 NCI_CGAP_Lu30 Mus musculus cdna clone IMAGE:4016413 5',
            mRNA sequence.
ACCESSION  BF141567
VERSION    BF141567.1 GI:10980607
KEYWORDS   EST.
SOURCE     Mus musculus (house mouse)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (Bases 1 to 1079)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: rgs@bbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM9263 row: 1 column: 14
High quality sequence stop: 663.
FEATURES
Location/Qualifiers
1..1079
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:4016413"
/tissue_type="tumor, metastatic to mammary"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Lu30"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; transgenic model WNT-1, expression driven by
MTMV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
dT. Library constructed by Life Technologies.
Investigator providing samples: Gilbert Smith, NIH"
BASE COUNT      312 a  267 c  263 g  237 t
ORIGIN

Query Match      39.2%; Score 497.6; DB 10; Length 1079;
Best Local Similarity 84.1%; Pred. No. 6.1e-139;
Matches 621; Conservative 0; Mismatches 109; Indels 8; Gaps 5;

QY 256 AAGAAGCAGGTTCCAGGCTGTGGTGTACTCAGCATGGCTAGTTGGAGTGTAGC 315
DB 20 AAGAAGGAAGATCCAGGCGAGTGTGTACTGTCAACATGGTCTCTTTGGGATGTAGC 79

QY 316 AACTGGATTTCCAACTGCCCAACAATAGCTGGCTTCATTCTGGCAGATGCTGGTTT 375
DB 80 AACTGGATTTCCAACTGCCCAACAACAGCTGGTGTATTTATTTGGCAGATGCAAGTTT 139

QY 376 GACGTGTGGATGGGGAACAGCAGGGGAAACGCC-TGGTCTCGAAAACACAGACACTCTC 434
DB 140 GATGTGTGGATGGGAAACAGCAGAGGAAACACCGTGGTCTCGAAGCAACAGCCCTCTC 199

QY 435 CATAGACCAAGATCAGTTCGGGCTTTTCAGTTATGATGAGATGGCTAGGTTTGACCTTCC 494
DB 200 CATGATCAAGATGAGTTCGGGCTTTTCAGTTATGATGAAATGGCTAGGTTTGACCTTCC 259

QY 495 TCGAGTGATAAACTTTATTTTGCAGAAAACGGGCCAGGAAAAGATCTATTATGTCGGCTA 554
DB 260 AGCTGTGATAAACTTTATCTTACAGAAAACGGGCCAGGAAAAGTCTATTATGTCGGCTA 319

QY 555 TTCACAGGGCACCACTGGGCTTTATTCATTTTCCACCATGCGAGCTGGCTCAGAA 614
DB 320 CTCACAGGGCACCACTGGGCTTTATTCGATTTTCCATGCGAGAGTGTAGCTCATAA 379

QY 615 AATCAAAATGATTTTGTCTTTAGCACCCTATAGCCACTGTTTAAGCATGCAAAAAGCCCGG 674
DB 380 AATCAAAATGATTTTGGCCCTTAGCTCTATAGCCACTGTTTAATATGCAAGAGTCTCGG 439

QY 675 GACCAAAATTTTGTCTGCGCAGATATGATCAAGGATTTGTTGGCAAAAAGAAATT 734
DB 440 TACCAAAATTTCTGTGTCGCGCAGATATGATCAAGGATTTATTTGGCAGACAAGATT 499

QY 735 TCTGTATCAGACCAAGATTTCTCAGACAACTTTGTTTATTTACCTTTGTGGCCAGGTGATTCT 794
DB 500 TTTATACCAAGACTAGATTTTTCAGGAGCTTTTATTTACCTTTGTGGCCAGATGATTCT 559

```

795 TGATCAGATTGTAGTAAATATCATGTTACTTCTCGGTGGATTCACCAACCAATATGAA 854
 Db CGACCAATCTGCAGCAACATCATCTTACTCTCGGAGGTTTAAACCAACCAATATGAA 619
 QY 855 CATGAGCGGAGCA--AGTGATATGCTGCCACACTCTTGTGGAACATCTGTG--CAA 909
 Db CATGAGCGAGCAACATGCTCTGTGCGCATTAACCTGCTGGAACGCTTGTGTCGCGA 679
 QY 910 AATATTCTACACTGGAGCAGG--CAGTGAATCTTGGTG--AACTCCGGCATTGACTGGG 967
 Db ACCATTCTCCACTGGAGCAGGCGGTGAATCTGGGCGACCTTCGTGCGCTTGTGACTGTG 739
 QY 968 GGAGTGAGACCAAAATC 985
 Db 740 GGCAAGCGAACCAAGC 757

RESULT 7
 AK088659
 LOCUS
 DEFINITION
 Mus musculus 2 days neonate thymus thymic cells cDNA, RIKEN full-length enriched library, clone:E430023A04 product:lysosomal acid lipase 1, full insert sequence.
 AK088659
 VERSION
 AK088659.1 GI:26353707
 KEYWORDS
 HTC; CAP trapper.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 99279253
 PUBMED
 10349636

2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 20499374
 PUBMED
 11042159

3
 Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipipillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 PUBMED
 11076861

4
 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamataka, I., Saito, T., Okazaki, Y., Gojohori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsumoto, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Stauble, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldairelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,

Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. P., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Wittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S. and Hayashizaki, Y.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)
 21085660
 PUBMED
 11217851

5
 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6 (bases 1 to 2138)

ADACHI, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirakawa, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, K., Ohno, M., Sakazume, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
 Direct Submission

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.

Please visit our web site for further details.
 URL: http://genome.gsc.riken.go.jp/
 Location/Qualifiers
 1. .2138
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="NOD"
 /db_xref="FANTOM DB:E430023A04"
 /db_xref="taxon:10090"
 /clone="E430023A04"
 /cell_type="thymic cells"
 /tissue_type="thymus"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="2 days neonate"
 66. .1259
 /note="unnamed protein product: lysosomal acid lipase 1 (WGD|GI:96789, GI|NM_021460, evidence: BLASTN, 99%, match=2084)
 /codon_start=1
 /protein_id="BAC40484.1"
 /db_xref="GI:26353708"
 /translation="MQLGLVFTVFTIGLLSRVPTGTVSAVDPEVNNVTEIMRWGY PEERSVLFGDGYILSIHRIHPRGNKFGKGRFPVYQLGHLLADSSWNVDNSSL GLLQADGFDVWNGSRNRTGKLSVQDFWAFSDFEDMADLPASINYLINK TQEQIYYVGHSGGTGTFIAFSPQMPFLAKIMFLVPLVSLNFAFGPLLQGLRPL DPLKDMFQKQFLPQSAAMLKLSHVCTHVMKLCANVFLCGLGFRNLMNRVD

FEATURES

source

CDS

/db_xref="FANTOM_DB:2310067K20"
 /db_xref="MGI:1898392"
 /db_xref="taxon:10090"
 /clone="2310067K20"
 /sex="male"
 /tissue_type="tongue"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"
 53..1240
 CDS
 /note="unnamed protein product; TRIACYLGLYCEROL LIPASE, LINGUAL PRECURSOR (EC 3.1.1.3) (LINGUAL LIPASE) homolog [Rattus norvegicus] (SWISSPROT:P04634, evidence: FASTV, 88.4%ID, 100%length, match=1184) putative"
 /codon_start=1
 /protein_id="BAB26697.1"
 /db_xref="MGI:12845298"
 /db_xref="MGI:1914967"
 /translation="MWLLVTVGLVAFGGAGHGLFGKLPKPNPEANNVVSOMITWYGYPSSEYEVVTDGVLGVIPYGVKNSNENIKRPVAYLOHGLVASATNWTNLPNNSLAFILADAGVDMVLGNSRGNTWRSKNTVYSPDSVEFWAFSPDEMAKYDLDPATIDFIVOKTGEKLIHVSHSQTTIGFIAPSTNPAKAKIKRFYALAPVATVKYTESPPFKLSLPKFLKLVIFGNKMFENYLDQFLGTEVCSRELLDLCSNALFIFCGFDKKNLNSRFDVYLGNHPAGTSDQLEFWAQLAKSGKLQAYNWSPLQNLMLHYNKTPPPYDYVSAMTVPYAVNGGHDLADPDQVAMLLPKLPNLLYHKEILPYNLHDFIWDAPQEVYNEIVTMVAED"
 BASE COUNT 396 a 290 c 286 g 378 t
 ORIGIN

Query Match 36.4%; Score 461.8; DB 11; Length 1350;

Best Local Similarity 63.8%; Pred. No. 5e-128;

Matches 700; Conservative 0; Mismatches 397; Indels 0; Gaps 0;

QY 125 ACCCAGAGCATTCATGATATTAGTGAATCATCAACATCAAGGCTATCCCTGTGAGG 184
 DB 132 ACCTGAGCAACATGATGTTAGTCAGATGATACTTACTGGGATATCCAGTGAGG 191
 QY 185 AATATGAAGTCGCAACTGAGATGGGTATATCTTTCTGTTAACAGGATCTCTCGAGGCC 244
 DB 192 AATATGAAGTTGTTACTGAAGATGGCTACATTCGGGGGTCTATAGAAATCTCTTATGGA 251
 QY 245 TAGTGCAACTAGACAGAGGTCACAGCGCTGTGTGTTACTGAGCATGCGCTAGTTG 304
 DB 252 AGAAAAATTCGAGAAATTCGGAAGAGACCTGTGCATATTTGCAAGATGTTGGTTG 311
 QY 305 GAGGTGCTAGCAACTGATTTCAACCTGCCAACATAGCCTGGCTTCATTCTGGCAG 364
 DB 312 CATCAGCCCAAACTGGATTACAAATCTGCCAAACACAGCGCTGGCTTCATTCTAGCAG 371
 QY 365 ATGCTGTTTGTAGGTGTGATGGGACAGAGGGGAAACGCTGGTCTCGAACAACA 424
 DB 372 ATGCTGCTATGATGTGTGCTGGGGAACAGTCCAGGGAATACATGGTCCCGAAAAATG 431
 QY 425 AGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTTCAGTTATGATGAGATGGCTAGT 484
 DB 432 TATACTATTACCAGACTCAGTTGAAATCTGGGCTTTTCAGTTATGATGAAATGGCTAAT 491
 QY 485 TTGACCTCTCGAGTGATATAAATCTTATTTTTCAGAAAAACGGCCAGGAAAGATCTATT 544
 DB 492 ATGACCTTCCAGCCACCATAGACTTCATTGTACAGAAAACTGGACAAGAGATACACT 551
 QY 545 ATGTCGCTATTACAGGGCACCACCATGGCTTTATGCTATTTCCACCATGCCAGAGC 604
 DB 552 ATGTTGTCTACTCTCAGGGCAACCACTATCGTTTATTTGCTTTTCTACCAATCTGCTC 611
 QY 605 TGGCTCAGAAAAATCAAAATGATTTTGTCTTTAGCAACCATAGCCACTGTTAGCATGCAA 664
 DB 612 TGGCTAAAAAATCAAGAGGTTTATGCAATGATGCTTCCAGTTGCTACTGTGAAGATACAG 671
 QY 665 AAGCCCGGGACCAAAATTTTGTGTCGCCAGATGATGATCAAGGGATGTTTGGCA 724
 DB 672 AAGTCCCTTTAAAAAGATTTCATCTTATCTTAAGTTTCTTCTCAAGGGTATATTGGTA 731

QY 725 AAAAAAGATTTCTGTATCAGACAGAGATTTCTCAGACAACTTTTATTTACCTTTTGGGC 784
 DB 732 AAAAAATGTTTCATGCCCCCAACTACTTAGATCAATTTCTTGGTAGGAGTGTGCTCAC 791
 QY 785 AGGTGATTTCTGTATCAGATTTGTAGTAATATCATGTATTCTTCTGGGTGATTCAACACA 844
 DB 792 GGGAGCTGTAGATCTTCTCTGAGCAACGCTTTATTATCTTCTGTGATTGACAAGA 851
 QY 845 ACAATATGAACATGAGCGGAGCAAGTGTATATGCTGCCACACTCTTCTGCTGGAACATCTG 904
 DB 852 AAAAACTTAATGTGAGTCGCTTTGATGTGTATCTAGGCAATAATCCAGGAGCAATCTA 911
 QY 905 TGCAAAATATTCTACACTGAGCAGGAGTGAATTTCTGTGAATCTCCGGGCAATTTGACT 964
 DB 912 CTCAAGACCTTTTCCACTGGCACAGCTTCTAAATCTGGAAGCTTCAAGCCTATAACT 971
 QY 965 GGGGAGTGAGACAAATAATCTGAAAAATGCAATCAGCAACTCTCTGTAAGGTACAGAG 1024
 DB 972 GGGGAGTCCATTTACAGAACATGTTTACTACATAACAGAAACGCTCCCTACTATGATG 1031
 QY 1025 TCAGAGATATGACGTTCCCTTACAGCAATGTGGACAGAGGTCAAGGACTGGCTTTCAAATC 1084
 DB 1032 TGTGAGCCATGACCGTGCCAAATTCAGTGTGGACGCGTGGCCATGACATCTTGGCTGATC 1091
 QY 1085 CAGAAGACGTGAAATATGCTCTCTCTGAGGTGACCAACCTCATCTACCATAGAAATTC 1144
 DB 1092 CCCAAGATGTGCAATGTCTTCCCAAACTCCCAACCTTCTGTACCATAGGAGATTC 1151
 QY 1145 CTGAATGGCTCAGCTGATTTTCATCTGGGGTTTGGATGCTCTCCTACCGTATGTACAATG 1204
 DB 1152 TTCCCTACAATCACTGGATTCATCTGGCGATGATGGCGCTCAAGAGTTTACAATG 1211
 QY 1205 AAATCATCCATCTGATG 1221
 DB 1212 AGATAGTTACCATGATG 1228

RESULT 10

AK010116

LOCUS

DEFINITION

AK010116

Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310068005 product:TRIACYLGLYCEROL LIPASE, LINGUAL PRECURSOR (BC 3.1.1.3) (LINGUAL LIPASE) homolog [Rattus norvegicus], full insert sequence.

ACCESSION

AK010116

VERSION

AK010116.1 GI:12845334

KEYWORDS

HTC; CAP trapper.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

PUBMED

10349636

REFERENCE

2

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

PUBMED

20499374

REFERENCE

3

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuami, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.


```

source
1. .1349
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:2310011G18"
/db_xref="MGI:1895751"
/db_xref="taxon:10090"
/clone="2310011G18"
/sex="male"
/tissue_type="tongue"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
53. 1240
/note="unnamed protein product; TRIACYLGLYCEROL LIPASE, LINGUAL PRECURSOR (EC 3.1.1.3) (LINGUAL LIPASE) homolog [Rattus norvegicus] (SWISSPROT|P04634, evidence: FASTA, 88.4%ID, 100%length, match=1184) putative"
/codon_start=1
/protein_id="BAB26201.1"
/db_xref="GI:12844012"
/db_xref="MGI:1914967"
/translation="MWLLLVSVLSAFGGHGLFKGLPKNPENNNVSMITWVGP SEEVETDGYILGVNRPYKKNSENIGRPVAYLQHLIASATNMTLNENSLA FILADAGVWILGSRNWTGRKNVYSPDSVEFWAFSDMAKYDIDFIDIVOKT GQEKHYVGHSGQTIGFIAFSTNPALAKIKRYFALAPVATVYTESPEPKISLIPK FLKVIYFONKMFPHNYLDQFLGTVEVCSRELLDLCSNALFIFCGPKNLYNRSFGV YLGHNPACTSDQLFHWALAKSKLQAYNWGSPLQNMHLNOKTPPYVDVSATVPI AVWNGHDIILADPDQVAMLLPKLPNLYHKELLPNHLDPIWMDAPQEVNIVTM AED"
BASE COUNT      395 a 290 c 284 g 380 t
ORIGIN
Query Match      36.3%; Score 460.2; DB 11; Length 1349;
Best Local Similarity 63.7%; Pred No. 1.5e-127;
Matches 699; Conservative 0; Mismatches 398; Indels 0; Gaps 0;

QY 125 ACCGAGAGCATTCATGAATATTAGTGAATCATCCAAATCAAGGCTATCCCTGTGAGG 184
DB 132 ACCCTGAAGCAACATGAATGTTAGTCAGATGATAAATCTTCTGGGGATATCCAAAGTCAGG 191
QY 185 AATATGAAGTCGCACTCAAGATGGTATATCCCTTCTGTGTAACAGGATTCCTCGAGGCC 244
DB 192 AATATGAAGTGTGTTACTGAAGATGCTACATCTTGGGGGCTCTAGAGATTCCTTATGGGA 251
QY 245 TAGTGCAACCTAAGAGACAGATTCAGGCTGTGGTGTACTGTCAGCATGCGCTAGTTG 304
DB 252 AGAAAAATTCGAGAATATCGGCAAGACCTGTGGCATATTTGCGCATGTTTGTATTG 311
QY 305 GAGGTGTAGCACTGGATTTCCAACTGCGCAACATAGCTGGGCTTCTATCTGGCAG 364
DB 312 CATCAGCACAACTGGGATTAACAATCTGCCAAACACAGCCTGGCCTTCTATCTAGCAG 371
QY 365 ATGCTGGTGTTCAGCTGTGGATGGGGAACAGCAGGGGAAACGCTGGTCTCGAAAAACA 424
DB 372 ATGCTGGCTATGATGTGTGGCTGGGGAACAGCTCGAGGAATACATGTCTCCGGAAAAATG 431
QY 425 AGACACTCTCCATAGACCAAGATAGTGTCTGGGCTTTTCAGTTATGATGAGATGCTAGGT 484
DB 432 TATACTATTCCAGCACTCAGTTGAATTTCTGGGCTTTTCAGCTTTTCATGAAATGCTAAAT 491
QY 485 TTGACCTTCTCTGCTGATTAATTTATTTTTCAGAAAACGGCGCAGGAAGATCTATT 544
DB 492 ATGACTTTCAGCCACCATAGACTTCATGTGACAGAAACCTGGACAGAGAGATACACT 551
QY 545 ATGTGCGGTATTACAGGGGACCAACCATGGGCTTTTATGCAATTTTCCACCATGCGCAGC 604
DB 552 ATGTTGGTCACTCTCAGGSCACCACTATCGGTTTTATTTGCTTTTCTTACCAATCCTGTC 611
QY 605 TGGCTCAGAAATCAAAATGATTTTGTCTTTAGCACCCATAGCCACTGTTAAGCATGCAA 664
DB 612 TGGCTAAAAAATCAAGAGGTTTTATGATGATAGTCTCCAGTTGCTACTGTGAAGTATACAG 671

```

665 AAAGCCCCGGGACCAAAATTTTGTGCTGCCAGATATGATGATCAAGGGATTTGTTGGCA 724

672 AAAGTCCCTTTAAAAAGATTTCACTTATTTCTTCAAGGTGATTAATTTGGTA 731

725 AAAAGAAATTTCTGTATCAGACAGATTTCTCAGACAACTTGTATTATTTACCTTTTGTGGCC 784

732 ACAAATGTTTCATGCCCACTACTTAGATCAATTTCTTGGTACGGAAGTGTCTCAC 791

785 AGGTGATTTCTGATCAGATTTGTAGTAATATCATGTTACTTCTGGGTGATTTCAACACCA 844

792 GCGAGCTGTAGATCTTCTCTGACGACACGCTTTATTCATCTTCTGTGATTTGCAAGA 851

845 ACAATATGAACATGAGCGGAGCAAGTGTATATGCTGCCACACTCTTGTGGAACATCTG 904

852 ACAAATTAATGTAGTCTGCTTTGTGTATCTAGGGAATATTCAGCAGGAAATCTTA 911

905 TGCAAAATATTTACATCTGAGCGGAGCAAGTGTGTAATCTTGTGTAATTCGCGGCAATTTGACT 964

912 CTCAAGACCTTTTCCACTGGGCAAGCTTGTCTGTAATCTGGGAAGCTTCAAGGCTATACT 971

965 GGGGAGTCAGACCAAAATCTGGAATAATCAATCAGCAACTCTCTGTAAAGGTACAGAG 1024

972 GGGGAAGTCCATTCAGAAACATGTTACTACTCAATCAGAAACGCTCCCTACTATGATG 1031

1025 TCAGAGATATGACGCTCCCTACAGCAATGTGACAGAGGTCAGGACTGGCTTTTCAAAATC 1084

1032 TGTACAGCCATGACCGTGCCAAATTCAGTGTGGAACGCTGGCCATGACATCTCTGGCTGATC 1091

1085 CAGAGACGTTGAAATGCTGCTCTCTGAGGTGACCAACTCATCTACCAATAGATATTC 1144

1092 CCAAGATGTGCAATGCTGCTTCCCAACTCCCAACCTCTGTACCAATAGGAGATTC 1151

1145 CTGAATGGGCTCACGTGGATTTTCATCTGGGTTTGGATGCTCTCCACCGTATGTACAATG 1204

1152 TTCCCTCAATCACCTGGATTCATCTGGCGCATGATGGGCTCAAGAGGTTTACAATG 1211

1205 AATCATCATCTGATG 1221

1212 AGATAGTTACCATGATG 1228

AK010139 1350 bp mRNA linear HTC 05-DEC-2002

Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310069P19 product:TRIACYLGLYCEROL LIPASE, LINGUAL PRECURSOR (EC 3.1.1.3) (LINGUAL LIPASE) homolog [Rattus norvegicus], full insert sequence.

AK010139 GI:12845371

HTC; CAP trapper.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

99279253

10349636

2

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, F., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

20499374

11042159

3

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

Qy	125	ACCCAGAGCATTCTATGATAATTAGTGAATCATCCAAATCAGGCTATCCCTGTGAGG	184
Db	126	ACCCCTGAAGACAAACATGAATCTTAGTCAGATGATAACTTACTGGGATATCCAAGTGGG	185
Qy	185	AATATGAAGTCGCAACTGAGATGGGTATATCTTTCTGTAAACAGGATTCCTCGAGGCC	244
Db	186	AATATGAAGTTGTTACTGTGAGATGGCTACATTTCTGGGGTCTATAGAATTCCTTATGGGA	245
Qy	245	TAGTGCAACCTAAGAAGACAGGTTCCAGGGCTGTGGTGTTCATCGCAGCATGCCCTAGTTG	304
Db	246	AGAAAAAATTCAGAGAATTCGGCAAGAGACCTGTGGCATATTTGCCAGCATGTTTGAATTG	305
Qy	305	GAGGTGCTAGCAACCTGGATTTCCAACTGCCCAACAAATAGCCTGGGCTTCATTCCTGGCAG	364
Db	306	CATCAGCCACAACCTGGGATTAACAATCTGCCAAACCAACGCTTGGCTTCATTCCTAGCAG	365
Qy	365	ATGCTGGTTTTGACCTGTGAGATGGGGAAACAGCAGGGGAAACGCTGTCTCGAAACACA	424
Db	366	ATGCTGGCTATGATGTGTGGCTGGGGAAACAGTCAGGGGAATACATGTCCTCCGGAAAAATG	425
Qy	425	AGNCACTCTCCATAGACCAAGATGAGTCTGGGGCTTTCAGTTATGATGAGATGGCTAGGT	484
Db	426	TATACTATTACCAAGACTCAGTTGAATCTCTGGGCTTTCAGCTTTGATGAAATGGCTAAAT	485
Qy	485	TTGACCTTCCTCAGTGTATAACTTTATTTTTCAGAAAAAGCGGCCAGGAAAAAGATCTATT	544
Db	486	ATGACCTTCCAGCCACCATAGACTTCATGTACAGAAAACTGGACCAAGAGAGATACACT	545
Qy	545	ATGTGGGCTATTCAAGGGCAACCAATGGGCTTTATTGCAATTTTCCACCATGCCAGAGC	604
Db	546	ATGTTGGTCACTCTCAGGGCAACCACTACGGTTTATTGGCTTTTCTACAACTCCTGCTC	605

Db 549 ATGTTGGTCACTCTCAGGCGCACCACTATCGGTTTTATTGCTTTTCTTACCAATCTGCTC 608
 Qy 605 TGCTCAGAAATCAAAATGATTTTCTTTAGCAGCCCATAGCCTGTTAGCATGCA 664
 Db 609 TGCTTAAATAATCAAGAGGTTTTATGTCATTAGCTCCAGTCTGCTAGTGAATATACAG 668
 Qy 665 AAAGCCCGGCGCAAAATTTTGTGCTGCGAGATGATGATCAAGGATTTGTTGGCA 724
 Db 669 AAAGTCCCTTAAAGATTTTCACTTATTCCTTAAGTTTCTTCTCAAGGTTGATTTGTA 728
 Qy 725 AAAAAGATTTCTGATACAGACAGATTTCTCAGAACTTGTATTTACCTTTGTTGGCC 784
 Db 729 ACAAATGTTTCATGCGCCCAACAATCTAGATCAATTTCTTGTACGGAAGTGTGCTCAC 788
 Qy 785 AGTGATTTCTGATCAGATTTGTAGTAATATCATGTTACTTCTGGTGGATTCACACCA 844
 Db 789 GGGAGCTGTAGATTTCTTCTGAGCAACGCTTTATTCATCTCTGTGATTTGACAGA 848
 Qy 845 ACAATATGAATGAGCGGAGCAAGTGTATGCTGCCACACTCTTTGCTGGAAATCTG 904
 Db 849 AAAAATTAATGTGAGTGGCTTTGATGTATCTAGGGCATATCCAGCAGGAACATCTA 908
 Qy 905 TCGAATATTTACACTGAGCGAGCAGCTGATTTCTGTGAATCTCGGCACTTTGACT 964
 Db 909 CTCAGACCTTTTCCACTGGGACAGCTTCTTAATTTGGAAAGCTTCAAGCCTTAAT 968
 Qy 965 GGGGAGTGAGACCAAAATCTGMAAAATGCAATCAGCAACTCTCTGTAAGTACAGAG 1024
 Db 969 GGGGAAGTCCATTACAGAACATGTTACATACATCAGAAACGCTCCCTACTATGATG 1028
 Qy 1025 TCAGATATGACGGTCCCTACAGCATGTGGACAGGTCAGGACTGGCTTCAATC 1084
 Db 1029 TGTGAGCATGACCGTCCCAATTCAGTGTGGAAACGCTGGCCAGCATCTCGGCTGATC 1088
 Qy 1085 CAGAAGACGTGAAATGCTGCTCTCAGGTGACCAACCTCATCTACCAATGAAATTC 1144
 Db 1089 CCCAAGATGCGCAATGCTGCTCCCAACCTCCCAACCTTCTGTACCAATGAGATTC 1148
 Qy 1145 CTGAATGGGCTCAGTGATTTGATCTGGGGTTGGATGCTCTCCACCGTATGTAATG 1204
 Db 1149 TTCCCTACATCAGCTGATCTATCTGGGCGATGGATGCGCTCAAGAGGTTTCAATG 1208
 Qy 1205 AAATCATCCATCTGATG 1221
 Db 1209 AGATAGTTACCATGATG 1225

RESULT 15
 LOCUS AK009413
 DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310020222 product:TRACVLGLYCEROL LIPASE, LINGUAL PRECURSOR (EC 3.1.1.3) (LINGUAL LIPASE) homolog (Rattus norvegicus), full insert sequence.
 ACCESSION AK009413
 VERSION AK009413.1 GI:12844193
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)
 99279253
 10349636
 2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)
 JOURNAL

MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

20499374
 11042159

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanishi, K., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multiplexed sequencer
 Genome Res. 10 (11), 1757-1771 (2000)

20530913
 11076861

MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L. M., Stauber, P., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombauts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. P., Suzuki, H., Taya, K., Wang, K. H., Weitz, C., Wittaker, C., Wilming, L., Wysshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohetsuki, S. and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)

21085660
 11217851

MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

5 The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

6 (bases 1 to 1345)
 Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hirao, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, I., Tejima, Y., Taya, T., Yamamura, T., Yanushita, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

Direct Submission
 Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN) Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.
 CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 04:41:17 ; Search time 350,095 Seconds
(without alignments)
9784.725 Million cell updates/sec

Title: US-10-042-431-46

Perfect score: 1269

Sequence: 1 atgttgaaacccctgtcaag.....gacgggtgagccgtatg 1269

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq 19Jun03.*

- 1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
- 25: /SIDSL1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1269	100.0	1269	22	Human TANGO 294 OR
2	1269	100.0	2044	22	Human TANGO 294 CD
3	1269	100.0	2044	25	cDNA encoding huma
4	1267.4	99.9	2035	22	Human CG162 (or C5
5	1235.4	97.4	1384	22	Human CG162 (or C5
6	1122	88.4	1206	24	cDNA encoding huma
7	1122	88.4	1210	24	Human cDNA encodin
8	1104.8	87.1	1272	24	cDNA encoding huma

9	752	59.3	1116	24	ABK86570	Human cDNA encodin
10	527.8	41.6	2481	24	ABN97383	Gene #3891 used to
11	526.2	41.5	2493	22	AAH02906	Human shear stress
12	489.2	38.6	1378	14	AAQ42310	RG1 precursor. Or
13	470	37.0	1367	7	AAH06085	Sequence encoding
14	470	37.0	1367	7	AAH06085	Sequence encoding
15	470	37.0	1367	17	AAH58916	Human stomach cell
16	468.4	36.9	1695	22	AAH57420	Human gastric lipa
17	464.4	36.6	1336	6	AAH50385	Rat prelingual lip
18	456.8	36.0	1411	24	ABL57144	Kid goat pregastric
19	454.2	35.8	1194	24	AAH27803	Human lysosomal ac
20	454	35.8	1228	24	ABK12385	cDNA encoding huma
21	450.8	35.5	1284	24	ABK85978	DNA encoding huma
22	449.8	35.4	8324	24	ABL57145	yeast YE-1 express
23	448.4	35.3	1473	24	ABK12388	cDNA encoding huma
24	446.8	35.2	1185	24	AAH31193	Human triacylglyce
25	444.2	35.0	1528	15	AAQ68388	Canine gastric lip
26	444.2	35.0	1528	17	AAH58915	Dog gastric lipase
27	444.2	35.0	1531	15	AAQ68389	Canine gastric lip
28	443.6	35.0	1200	22	AAH78205	Nucleotide sequenc
29	424.8	33.5	1260	24	ABH71703	DNA encoding huma
30	419.8	33.1	1086	25	ACC44783	Human triacylglyce
31	414.2	32.6	1360	24	AAH47480	Human lysosomal ac
32	414.2	32.6	1360	25	ABX14880	Human lipase cDNA.
33	374.4	29.5	996	25	ACC48563	Human dithp secret
34	346.4	27.3	1138	24	ABQ82326	Human NOV1 encodin
35	332.8	26.2	1098	24	AAH27802	Human lysosomal ac
36	319.4	25.2	1143	24	ABK12386	cDNA encoding huma
37	315.4	24.9	828	24	AAH31194	Human triacylglyce
38	300.4	23.7	1377	24	ABH71702	DNA encoding huma
39	280	22.9	842	24	AAH31192	Human triacylglyce
40	286	22.5	699	22	AAF28679	Human protein HP03
41	286	22.5	1308	22	AAF28689	Human protein HP03
42	282.8	22.3	802	22	AAH7406	Human neuroblastom
43	267.4	21.1	18554	24	ABH55357	Genomic DNA encodi
44	257.8	20.3	617	21	AAA44349	Human secreted exp
45	244.6	19.3	1971	23	AAH91026	DNA encoding novel

ALIGNMENTS

RESULT 1
ID AAF45132 standard; cDNA; 1269 BP.
XX AAF45132;
XX AAF45132;
DT 30-MAR-2001 (first entry)
XX Human TANGO 294 ORF.

XX Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
XX central nervous system; focal brain disorder; bipolar affective disorder;
XX global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
XX senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
XX Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
XX neuropsychiatric; psychoactive substance use; anxiety; ss.
XX Homo sapiens.
XX WO200077239-A2.
XX 21-DEC-2000.
XX 24-MAY-2000; 2000WO-US14858.
XX 14-JUN-1999; 99US-0333159.
XX (MILL-) MILLENNIUM PHARM INC.
PI McCarthy SA, Fraser CC, Sharp JD, Barnes TM;

DR WPI: 2001-032313/04.
 DR P-PSDE; AAB66065.
 XX
 TT TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
 PT screening assays and diagnostic assays and for the treatment of
 PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
 PT disease -
 XX
 XX Claim 1; Fig 6; 359pp; English.
 XX
 CC The present invention relates to TANGO or INTERCEPT proteins and coding
 CC sequences (see AAF45131-F45136 and AAF45138-F45139 and AAB66031-B66057,
 CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
 CC sequences are useful for the treatment of neurological disorders such as
 CC central nervous system (CNS) disorders, CNS-related disorders, focal
 CC brain disorders, global-diffuse cerebral disorders and other
 CC neurological and cerebrovascular disorders. The CNS disorders include
 CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
 CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
 CC autonomic function disorders such as hypertension and sleep disorders,
 CC neuropsychiatric disorders, psychoactive substance use disorders,
 CC anxiety, and bipolar affective disorder.
 XX
 SQ Sequence 1269 BP; 358 A; 268 C; 309 G; 334 T; 0 other;
 Query Match 100.0%; Score 1269; DB 22; Length 1269;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGTTGGAAACCTTGTACAGACAGTGGTATGCTACACAGATGGAATGTGGCTTCTG 60
 DB 1 ATGTTGGAAACCTTGTACAGACAGTGGTATGCTACACAGATGGAATGTGGCTTCTG 60
 QY 61 ATTCTGGTGGCGTATATGTTCCAGAGAAATGTGAATTCAGTACATATGCAACATAAGCT 120
 DB 61 ATTCTGGTGGCGTATATGTTCCAGAGAAATGTGAATTCAGTACATATGCAACATAAGCT 120
 QY 121 GTGACCCAGAACATTCATGAATATATAGTGAATCATCCACATCAAGGCTATCCCTGT 180
 DB 121 GTGACCCAGAACATTCATGAATATATAGTGAATCATCCACATCAAGGCTATCCCTGT 180
 QY 181 GAGGAATATCAAGTCGCACTGAGATGGTATATCTTCTGTTAAACAGGATTCCTCGA 240
 DB 181 GAGGAATATCAAGTCGCACTGAGATGGTATATCTTCTGTTAAACAGGATTCCTCGA 240
 QY 241 GGCTAGTGCACCTCAAGAGACAGGTTCCAGGCTGTGGTGTACTGACAGATGGCCTA 300
 DB 241 GGCTAGTGCACCTCAAGAGACAGGTTCCAGGCTGTGGTGTACTGACAGATGGCCTA 300
 QY 301 GTTGGAGGTGCTAGCACTGGATTTCCAACTGCCCAACATAGCTGGGCTTCATCTG 360
 DB 301 GTTGGAGGTGCTAGCACTGGATTTCCAACTGCCCAACATAGCTGGGCTTCATCTG 360
 QY 361 GCAGATGCTGGTGTGAGTGGGAAACAGCAGGGGAAACGCTGTGTTCTCGAAA 420
 DB 361 GCAGATGCTGGTGTGAGTGGGAAACAGCAGGGGAAACGCTGTGTTCTCGAAA 420
 QY 421 CACAGACACTCTCCATAGACCAAGATGCTTGGGCTTTCAGTATGATGATGGCT 480
 DB 421 CACAGACACTCTCCATAGACCAAGATGCTTGGGCTTTCAGTATGATGATGGCT 480
 QY 481 AGGTTTGCACCTTCTGCACTGATAAATCTTTATTTGAGAAACGGGCCAGGAAAGATC 540
 DB 481 AGGTTTGCACCTTCTGCACTGATAAATCTTTATTTGAGAAACGGGCCAGGAAAGATC 540
 QY 541 TATTATGTGGCTATTACAGGGGACCAACATGGGCTTTATGATTTTCCACCATGCCA 600
 DB 541 TATTATGTGGCTATTACAGGGGACCAACATGGGCTTTATGATTTTCCACCATGCCA 600
 QY 601 GAGCTGGCTCAGAAAATCAAAATGTAATTTGCTTTAGCACCCATAGCCATGTTAAGCAT 660
 DB 601 GAGCTGGCTCAGAAAATCAAAATGTAATTTGCTTTAGCACCCATAGCCATGTTAAGCAT 660

QY 661 GCAAAAGCCCGGGACCAAAATTTTGTGTCGCAGATATGATGATCAAGGATTTGTT 720
 DB 661 GCAAAAGCCCGGGACCAAAATTTTGTGTCGCAGATATGATGATCAAGGATTTGTT 720
 QY 721 GGCACAAAAGAAATTTCTGTATCAGACCAAGATTTCTCAGACAACTTGTATTACCTTTGT 780
 DB 721 GGCACAAAAGAAATTTCTGTATCAGACCAAGATTTCTCAGACAACTTGTATTACCTTTGT 780
 QY 781 GSCCAGGTGATTTCTGATCAGATTTCTAGTATATATATCTTCTGGTGGATTTCAAC 840
 DB 781 GSCCAGGTGATTTCTGATCAGATTTCTAGTATATATATCTTCTGGTGGATTTCAAC 840
 QY 841 ACCAACAAATATGAACATGAGCCGAGCAAGTATATGCTGCCACACTCTTGTCTGGAACA 900
 DB 841 ACCAACAAATATGAACATGAGCCGAGCAAGTATATGCTGCCACACTCTTGTCTGGAACA 900
 QY 901 TCTGTGCAAAATTTCTACACTGGAGCCAGGAGTGAATTTCTGGTGAATTCGGGGCATTT 960
 DB 901 TCTGTGCAAAATTTCTACACTGGAGCCAGGAGTGAATTTCTGGTGAATTCGGGGCATTT 960
 QY 961 GACTGGGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCCAACTCTCTGAAGGTAC 1020
 DB 961 GACTGGGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCCAACTCTCTGAAGGTAC 1020
 QY 1021 AGAGTCAGATATGAGCGTCCCTCAGCAATGTGACAGGAGTCAAGACTGGCTTTCA 1080
 DB 1021 AGAGTCAGATATGAGCGTCCCTCAGCAATGTGACAGGAGTCAAGACTGGCTTTCA 1080
 QY 1081 AATCCAGAAGACGTGAAATGCTGCTCTCAGGTGACCAACCTCATCTACCATTAAGAAT 1140
 DB 1081 AATCCAGAAGACGTGAAATGCTGCTCTCAGGTGACCAACCTCATCTACCATTAAGAAT 1140
 QY 1141 ATTCTGAAATGGGCTCAGTGGATTTCACTGGGGTTTGGATGCTCTCTACCGTATGTAC 1200
 DB 1141 ATTCTGAAATGGGCTCAGTGGATTTCACTGGGGTTTGGATGCTCTCTACCGTATGTAC 1200
 QY 1201 AATCAATATCCATCTGATGACAGGAGGAGCAACCTTTCCAGGAGCGTGTGAG 1260
 DB 1201 AATCAATATCCATCTGATGACAGGAGGAGCAACCTTTCCAGGAGCGTGTGAG 1260
 QY 1261 GCCGTATTG 1269
 DB 1261 GCCGTATTG 1269
 RESULT 2
 AAF45131
 ID AAF45131 standard; cDNA; 2044 BP.
 XX
 AC AAF45131;
 XX
 DT 30-MAR-2001 (first entry)
 XX
 DE Human TANGO 294 cDNA.
 XX
 KW Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
 KW central nervous system; focal brain disorder; bipolar affective disorder;
 KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
 KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
 KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
 KW neuropsychiatric; psychoactive substance use; anxiety; ss.
 OS Homo sapiens.
 XX
 XX WO200077239-A2.
 XX
 PD 21-DEC-2000.
 XX
 PF 24-MAY-2000; 2000WO-US14858.
 XX
 PR 14-JUN-1999; 99US-0333159.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.

XX McCarthy SA, Fraser CC, Sharp JD, Barnes TM;
 XX WPI; 2001-032313/04.
 XX DR P-PSDB; AAB66065.
 XX
 XX TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
 PT screening assays and diagnostic assays and for the treatment of
 PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
 PT disease -
 XX
 XX Claim 1; Fig 6; 359pp; English.
 XX
 XX The present invention relates to TANGO or INTERCEPT proteins and coding
 CC sequences (see AAF451121-F451136 and AAF451138-F451139 and AAB66031-B66057,
 CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
 CC sequences are useful for the treatment of neurological disorders such as
 CC central nervous system (CNS) disorders, CNS-related disorders, focal
 CC brain disorders, global-diffuse cerebral disorders and other
 CC neurological and cerebrovascular disorders. The CNS disorders include
 CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
 CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
 CC autonomic function disorders such as hypertension and sleep disorders,
 CC neuropsychiatric disorders, psychoactive substance use disorders,
 CC anxiety, and bipolar affective disorder.
 XX
 XX Sequence 2044 BP; 618 A; 401 C; 460 G; 565 T; 0 other;
 XX
 XX Query Match 100.0%; Score 1269; DB 22; Length 2044;
 XX Best Local Similarity 100.0%; Pred. No. 0;
 XX Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 1 ATGTGGAAACCTTTGTCAAGACAGTGGATTGCTTCTCACAGAAATGGAAATGTGGCTTCTG 60
 Db 126 ATGTGGAAACCTTTGTCAAGACAGTGGATTGCTTCTCACAGAAATGGAAATGTGGCTTCTG 185
 QY 61 ATTCGTGGGGTATATGTTCCAGAGAAATGTGAATTCAGTACATATGCCAATCAAGCT 120
 Db 186 ATTCGTGGGGTATATGTTCCAGAGAAATGTGAATTCAGTACATATGCCAATCAAGCT 245
 QY 121 GTGACCCAGCAAGCAATTCATGAATATAGTGAATATCCAAATCAAGCTATCCCTGT 180
 Db 246 GTGACCCAGCAAGCAATTCATGAATATAGTGAATATCCAAATCAAGCTATCCCTGT 305
 QY 181 GAGGAATATGAGTGGCACTGAGATGGGTATATCTTTCTGTATACAGAAATTCCTCGA 240
 Db 306 GAGGAATATGAGTGGCACTGAGATGGGTATATCTTTCTGTATACAGAAATTCCTCGA 365
 QY 241 GGCCTAGTGCACCTTAAGACACAGCTTCCAGGCTGTGGTGTACTGACATGGCTA 300
 Db 366 GGCCTAGTGCACCTTAAGACACAGCTTCCAGGCTGTGGTGTACTGACATGGCTA 425
 QY 301 GTTGAGGTGTAGCACTGGAATTTCCAACTGCCCCCAATAGCCTGGGCTTCATTCTG 360
 Db 426 GTTGAGGTGTAGCACTGGAATTTCCAACTGCCCCCAATAGCCTGGGCTTCATTCTG 485
 QY 361 CGAGATGCTGGTGTGGTGTGGATGGGACACAGCGGGAACGCTGCTCGAAA 420
 Db 486 CGAGATGCTGGTGTGGTGTGGATGGGGAACACAGCGGGAACGCTGCTCGAAA 545
 QY 421 CACAAGACACTCTCCATGACCAAGATGAGTTCTGGGCTTTTCAGTTATGATGAGATGGCT 480
 Db 546 CACAAGACACTCTCCATGACCAAGATGAGTTCTGGGCTTTTCAGTTATGATGAGATGGCT 605
 QY 481 AGGTTGACCTTCTGCACTGATTAACCTTTATTTTGAGAAACCGGCGCAGAAAGATC 540
 Db 606 AGGTTGACCTTCTGCACTGATTAACCTTTATTTTGAGAAACCGGCGCAGAAAGATC 665
 QY 541 TATTATGCGGCTATTTCAGGGGACCAACCATGGGCTTTATGCAATTTTCCACCATGCCA 600
 Db 666 TATTATGCGGCTATTTCAGGGGACCAACCATGGGCTTTATGCAATTTTCCACCATGCCA 725
 QY 601 GAGTGGCTCAGAAAATCAAAATGATTTTGTGCTTTTACGCCCATAGCACTGTTAAGCAT 660

Db 726 GAGTGGCTCAGAAAATCAAAATGATTTTCTTAGCACCATAGCCACTGTTAAGCAT 785
 QY 661 GCAAAAAGCCCGGAGCAAAATTTTGTGTGTCAGATATGATGATCAAGGATTTT 720
 Db 786 GCAAAAAGCCCGGAGCAAAATTTTGTGTGTCAGATATGATGATCAAGGATTTT 845
 QY 721 GCAAAAAGCAATTTTCTGTATCAGACCAAGATTTCTCAGCAAACTTTGTTATTACCTTTGT 780
 Db 846 GCAAAAAGCAATTTTCTGTATCAGACCAAGATTTCTCAGCAAACTTTGTTATTACCTTTGT 905
 QY 781 GGCACAGGTGATTTCTGATCAGATTTGTAGTAAATATATGTTACTTCTGGGTGGATTCAC 840
 Db 906 GGCACAGGTGATTTCTGATCAGATTTGTAGTAAATATATGTTACTTCTGGGTGGATTCAC 965
 QY 841 ACCAACAATATGAAATCAGCGGAGCAAGTGTATATGTCGCCACACACTTTGCTGGAACA 900
 Db 966 ACCAACAATATGAAATCAGCGGAGCAAGTGTATATGTCGCCACACACTTTGCTGGAACA 1025
 QY 901 TCTGTGCAAAATATTTCTACCTGAGCGCAGGAGTAAATCTGTGGAATCTCGGSCATTT 960
 Db 1026 TCTGTGCAAAATATTTCTACCTGAGCGCAGGAGTAAATCTGTGGAATCTCGGSCATTT 1085
 QY 961 GACTGGGGAGTGCAGACCAAAATCTGGAATAATGCAATCAGCAACTCTCTGTAAGTAC 1020
 Db 1086 GACTGGGGAGTGCAGACCAAAATCTGGAATAATGCAATCAGCAACTCTCTGTAAGTAC 1145
 QY 1021 AGAGTCAGAGATATGACGGTCCCTACAGCAATGTGGACAGAGGTGAGGCTGCTTTCA 1080
 Db 1146 AGAGTCAGAGATATGACGGTCCCTACAGCAATGTGGACAGAGGTGAGGCTGCTTTCA 1205
 QY 1081 AATCCAGAAAGAGTGAATAATGCTCTCTGAGGTGACCAACCTCATCTACCAATGAAT 1140
 Db 1206 AATCCAGAAAGAGTGAATAATGCTCTCTGAGGTGACCAACCTCATCTACCAATGAAT 1265
 QY 1141 ATTCTGTAATGGGCTCACGTGATTTTCATCTGGGGTTTGGATGCTCTCACCGTATGAC 1200
 Db 1266 ATTCTGTAATGGGCTCACGTGATTTTCATCTGGGGTTTGGATGCTCTCACCGTATGAC 1325
 QY 1201 AATGAATATCATCTGATGATGACAGGAGGAGACCACTTTCCAGGAGCGGTGTGAG 1260
 Db 1326 AATGAATATCATCTGATGATGACAGGAGGAGACCACTTTCCAGGAGCGGTGTGAG 1385
 QY 1261 GCGGTATTG 1269
 Db 1386 GCGGTATTG 1394
 XX
 XX RESULT 3
 XX ABX94106
 XX ID ABX94106 standard; cDNA; 2044 BP.
 XX AC ABX94106;
 XX XX
 XX DT 03-JUN-2003 (first entry)
 XX XX
 XX DE cDNA encoding human TANGO 294.
 XX XX
 XX KW Human; TANGO 294; INTERCEPT; cellular process; tissue typing;
 KW forensic biology; cellular protease activity; cell interaction;
 KW development; blood disorder; haematopoietic cell-related disorder;
 KW growth; cell proliferation; cell differentiation; gamma delta T-cell;
 KW immune system; metabolic disorder; homeostatic disorder; anaemia;
 KW developmental bone disorder; osteoporosis; bacterial infection;
 KW thrombocytopaenia; renal failure; liver disease; cystic fibrosis;
 KW obesity; cancer; bronchitis; asthma; emphysema; pulmonary oedema;
 KW respiratory distress syndrome; osteopathic; antibacterial; antianemic;
 KW thrombolytic; nephrotropic; anitobesity; hepatotropic; cytostatic;
 KW antiinflammatory; antiasthmatic; gene; ss.
 XX
 XX OS Homo sapiens.
 XX XX
 XX PH Key Location/Qualifiers

CDS 126..1397
 FT /cag= a
 FT /product= "TANGO 294"
 FT /note= "The ORF given as SEQ ID No:46 is specifically
 FT claimed in Claim 2"
 XX
 PN US2002182675-A1.
 XX
 PD 05-DEC-2002.
 XX
 PF 25-OCT-2001; 2001US-0042431.
 XX
 PR 14-JUN-1999; 99US-0333159.
 PR 24-MAY-2000; 2000US-0578063.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI McCarthy SA, Fraser CC, Sharp JD, Barnes TM;
 XX
 DR WPI; 2003-328617/31.
 DR P-PSDB; ABU08369.
 XX
 PT New TANGO and INTERCEPT proteins, useful as modulating agents in
 PT regulating a variety of cellular processes, in chromosome mapping, in
 PT tissue typing, and in forensic biology -
 XX
 PS Claim 2; Fig 6A-6C; 232pp; English.
 XX
 CC The present invention relates to the isolation of novel TANGO or
 CC INTERCEPT proteins, and the polynucleotide sequences encoding them.
 CC The invention discloses sequences for human TANGO 202, TANGO 234,
 CC TANGO 265, TANGO 273, TANGO 286, TANGO 294 and INTERCEPT 296. The
 CC sequences for murine TANGO 202 and TANGO 273 are also provided. The
 CC TANGO polypeptides and the polynucleotide sequences encoding them are
 CC useful as modulating agents in regulating a variety of cellular
 CC processes. The polynucleotide sequences are useful as primers or
 CC hybridisation probes for the detection of nucleic acids encoding
 CC TANGO polypeptides, chromosome mapping, tissue typing, and in
 CC forensic biology. TANGO 202 can be used to diagnose, prevent or
 CC treat disorders relating to aberrant cellular protease activity,
 CC inappropriate interaction of cells with mediators, inappropriate
 CC development, and blood and haematopoietic cell-related disorders.
 CC TANGO 234 can be used to modulate growth, proliferation, survival,
 CC differentiation, and activity of gamma delta T-cells. TANGO 265 can
 CC be used to prevent, diagnose and treat disorders characterised by
 CC aberrant organisation or development of a tissue or organ, and for
 CC modulating differentiation of cells of the immune system. TANGO 273
 CC is useful for diagnosing, treating or preventing e.g. metabolic,
 CC homeostatic and developmental bone disorders (e.g. osteoporosis),
 CC and bacterial infection. TANGO 286 is useful for treating or
 CC preventing e.g. anaemia, thrombocytopaenia, renal failure or
 CC liver disease. TANGO 294 is useful for treating or preventing
 CC e.g. cystic fibrosis or obesity. INTERCEPT 296 is useful for
 CC diagnosing, treating or preventing e.g. cancers, bronchitis, cystic
 CC fibrosis, asthma, emphysema, pulmonary oedema, or adult and infant
 CC respiratory distress syndromes. The present sequence encodes
 CC human TANGO 294.
 XX
 XX Sequence 2044 BP; 618 A; 401 C; 460 G; 565 T; 0 other;
 Query Match 100.0%; Score 1269; DB 25; Length 2044;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGTTGGAACTTGTCCAGACATGATGTTCTCCACAGAAATGGAAATGGCTTCG 60
 DB |||||
 QY 126 ATGTTGGAACTTGTCCAGACATGATGTTCTCCACAGAAATGGAAATGGCTTCG 185
 DB |||||
 QY 61 ATTCTGGGGGTATATGTTCCAGAGAAATGTAATTCAGTACATATGCCAACTAAAGCT 120
 DB |||||
 QY 186 ATTCTGGGGGTATATGTTCCAGAGAAATGTAATTCAGTACATATGCCAACTAAAGCT 245
 DB |||||
 QY 121 GTGAGCCCAAGCAATTCATGATATAGTGAATCATCAACATCAAGGCTATCCCTGT 180
 DB |||||

246 GTGAGCCCAAGCAATTCATGATATAGTGAATCATCAACATCAAGGCTATCCCTGT 305
 QY |||||
 181 GAGGAATATGAGTCCGCACTGAAGATGGTATATCTTCTCTGTTAAAGGATTCCTCGA 240
 DB |||||
 306 GAGGAATATGAGTCCGCACTGAAGATGGTATATCTTCTCTGTTAAAGGATTCCTCGA 365
 DB |||||
 241 GGCCTAGTGCACCACTAAGAAAGACAGAGTTCAGAGCCCTGTGTGTTACTGACAGCATGCGCTA 300
 QY |||||
 366 GGCCTAGTGCACCACTAAGAAAGACAGAGTTCAGAGCCCTGTGTGTTACTGACAGCATGCGCTA 425
 DB |||||
 301 GTTGGAGGTGCTAGCAAACTGGATTTCCAACTCCGCCCAACATAGCTGGCTGGCTTCATCTG 360
 QY |||||
 426 GTTGGAGGTGCTAGCAAACTGGATTTCCAACTCCGCCCAACATAGCTGGCTGGCTTCATCTG 485
 DB |||||
 361 GCAGATGCTGTTTGGACGTGTGATGGGAAACAGAGGGGAAACGCTGTGTTCTCGAAAA 420
 QY |||||
 486 GCAGATGCTGTTTGGACGTGTGATGGGAAACAGAGGGGAAACGCTGTGTTCTCGAAAA 545
 DB |||||
 421 CACAAGACATCTCCATAGACCAAGATGAGTTCCTGGGCTTTCAGTTATGATGATGATGGCT 480
 QY |||||
 546 CACAAGACATCTCCATAGACCAAGATGAGTTCCTGGGCTTTCAGTTATGATGATGATGGCT 605
 DB |||||
 481 AGGTTTCACTCTCTGCAAGTATTAATTTTTCAGAAAAACGCGCCAGGAAAGATC 540
 QY |||||
 606 AGGTTTCACTCTCTGCAAGTATTAATTTTTCAGAAAAACGCGCCAGGAAAGATC 665
 DB |||||
 541 TATTATGTCGGCTATTCCAGGGGCAACCATGGGCTTTATTCATTTTCCACATGCCA 600
 QY |||||
 666 TATTATGTCGGCTATTCCAGGGGCAACCATGGGCTTTATTCATTTTCCACATGCCA 725
 DB |||||
 601 GAGCTGCTCAGAAATTAATTAATTTTTCAGTATGATGATGATGATGATGATGATGATGATGAT 660
 QY |||||
 726 GAGCTGCTCAGAAATTAATTAATTTTTCAGTATGATGATGATGATGATGATGATGATGATGAT 785
 DB |||||
 661 GCAAAAAACCCCGGAGCAAAATTTTCTGCTGCCAGATATGATGATGATGATGATGATGATGATGAT 720
 QY |||||
 786 GCAAAAAACCCCGGAGCAAAATTTTCTGCTGCCAGATATGATGATGATGATGATGATGATGATGAT 845
 DB |||||
 721 GCGAAAAAAGAAATTTCTGATCAGACAGATTTCTCAGACAACTGTTATTTACCTTTGT 780
 QY |||||
 846 GCGAAAAAAGAAATTTCTGATCAGACAGATTTCTCAGACAACTGTTATTTACCTTTGT 905
 DB |||||
 781 GCGAGGTGATTTCTGATCAGATTTGATGATTAATCATGTTTCTTGGTGGATTCAC 840
 QY |||||
 906 GCGAGGTGATTTCTGATCAGATTTGATGATTAATCATGTTTCTTGGTGGATTCAC 965
 DB |||||
 841 ACCAACATATGAACATGAGCCGAGCAAGTATATGCTGCCCACTCTTGTCTGGAACA 900
 QY |||||
 966 ACCAACATATGAACATGAGCCGAGCAAGTATATGCTGCCCACTCTTGTCTGGAACA 1025
 DB |||||
 901 TCTGTGCAAAATATTTCTACCTGAGCCAGCAGTGAATTTCTGTGAATCTCCGGGCAATT 960
 QY |||||
 1026 TCTGTGCAAAATATTTCTACCTGAGCCAGCAGTGAATTTCTGTGAATCTCCGGGCAATT 1085
 DB |||||
 961 GACTGGGGGAGTGAGACCAAAATCTGGAAAAATGCAATCAGCCAACTCTCTGTAAGTAC 1020
 QY |||||
 1086 GACTGGGGGAGTGAGACCAAAATCTGGAAAAATGCAATCAGCCAACTCTCTGTAAGTAC 1145
 DB |||||
 1021 AGAGTCAGAGATATGACGGTCCCTACAGCAATGAGCAGAGGTGAGAGGTGAGAGGTGAGAGGTGAG 1080
 QY |||||
 1146 AGAGTCAGAGATATGACGGTCCCTACAGCAATGAGCAGAGGTGAGAGGTGAGAGGTGAGAGGTGAG 1205
 DB |||||
 1081 AATCCAGAAAGACGTGAAATGCTCTCTGAGGTGACCAACCTCATCTACCAATGAAT 1140
 QY |||||
 1206 AATCCAGAAAGACGTGAAATGCTCTCTGAGGTGACCAACCTCATCTACCAATGAAT 1265
 DB |||||
 1141 ATTCTGTAATGGGCTCAGGTGGAATTCATCTGGGGTTTGGATGCTCTCCACCGTATGTAC 1200
 QY |||||
 1266 ATTCTGTAATGGGCTCAGGTGGAATTCATCTGGGGTTTGGATGCTCTCCACCGTATGTAC 1325
 DB |||||
 1201 AATGAATCATCCATCTGATCAGCAGGAGGAGCAACCTTTCCAGGAGCGGTGTGAG 1260
 QY |||||

Db 1326 AATGAATCATCTCTGATGAGCAGGAGGAGACCAACCTTCCAGGAGCGGTGTGAG 1385
 QY 1261 GCGGTATTG 1269
 Db 1386 GCGGTATTG 1394

RESULT 4

AAD19226
 ID AAD19226 standard; DNA; 2035 BP.

AC AAD19226;

DT 18-DEC-2001 (first entry)

XX

DE Human CG162 (or C59) lipase DNA #2.

XX Human; apolipoprotein; lipase; lipoprotein receptor; ALR; angina;
 KW cardiovascular disease; lipid metabolism; myocardial infarction;
 KW cerebral ischaemia; arterial thrombosis; thrombolytic; antilipemic;
 KW coronary artery thrombosis; cerebral artery thrombosis; stroke;
 KW intracardiac thrombosis; gene therapy; cardiovascular; vasodilator;
 KW neuroprotectant; cerebroprotective; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

PH 189..1418

FT /tag= a

FT /product= "Human CG162 (or C59) lipase protein"

FT 189..245

FT /tag= b

FT 246..1415

FT /tag= c

FT /product= "Human mature CG162 (or C59) lipase protein"

XX

PN WO200179446-A2.

XX

PD 25-OCT-2001.

XX

PF 16-APR-2001; 2001WO-US12529.

XX

PR 14-APR-2000; 2000US-197137P.

XX

PR 20-JUN-2000; 2000US-0598042.

XX

PR 03-AUG-2000; 2000US-0631451.

XX

PR 22-SEP-2000; 2000US-0667298.

XX

PR 17-NOV-2000; 2000US-0714936.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Ballinger DG, Loeb D, Montgomery JR, Tang TY, Zhou P, Goodrich R;
 PI Liu C, Asundi V, Zhao QA, Wehrman T, Drmanac RT, Ren F, Qian XB;
 PI Wang D;
 XX
 DR WPI: 2001-611724/70.
 DR P-PSDB; ABE11931.
 XX
 PT Nucleic acids encoding human apolipoproteins, lipases, and lipoprotein
 PT receptor polypeptides, useful for preventing diagnosing and treating
 PT lipid metabolism disorders, thrombosis and cardiovascular diseases -
 XX
 PS Claim 1; Page 183-185; 266pp; English.
 CC
 XX The invention relates to polynucleotides encoding proteins CG122, CG179,
 CC CG95, CG121, CG162, CG27, CG153 and CG168 which are related to proteins
 CC involved in lipid metabolism and cardiovascular disease such as human
 CC apolipoproteins, lipases and lipoprotein receptor proteins. These DNA
 CC and protein sequences are useful for treating or preventing disorders
 CC associated with apolipoproteins, lipases and lipoprotein receptor (ALR)
 CC expression and for treating lipid metabolism, cardiovascular diseases
 CC and thrombosis. Antibodies against these proteins are useful for
 CC determining the presence of or predisposition to a disease associated

CC with altered levels of these sequences. ALR polypeptides are also
 CC useful for identifying agents (agonists and antagonists) that bind to
 CC them and cells expressing ALR proteins are useful for identifying a
 CC therapeutic agent for use in treatment of a pathology related to
 CC aberrant expression or physiological interactions of this polypeptide.
 CC Vectors comprising these DNA and protein sequences are also useful for
 CC producing ALR proteins. The nucleic acids and polypeptides of the
 CC invention are also useful for the treatment of occlusive cardiovascular
 CC diseases, myocardial infarction, cerebral ischaemia, angina, arterial
 CC thrombosis, coronary artery thrombosis and cerebral artery thrombosis
 CC or intracardiac thrombosis and stroke. The nucleotides of the invention
 CC are used in gene therapy. The present sequence is human CG162 (or C59)
 CC lipase DNA.

XX
 SQ Sequence 2035 BP; 508 A; 395 C; 456 G; 576 T; 0 other;

Query Match 99.9%; Score 1267.4; DB 22; Length 2035;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1268; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGTTGGAAACCTTGTCAAGACAGTGGATTGTCTCACACAGATGCAATGGCTTCG 60
 Db |||||||
 147 ATGTTGGAAACCTTGTCAAGACAGTGGATTGTCTCACACAGATGCAATGGCTTCG 206

QY 61 ATTCTGGTGGCGTATATGTTCCAGAGAAATGGAATTCAGTACATATGCCAACTAAAGCT 120
 Db |||||||
 207 ATTCTGGTGGCGTATATGTTCCAGAGAAATGGAATTCAGTACATATGCCAACTAAAGCT 266

QY 121 GTGACCCAGACCATTCATGAATATAGTGAATCCACATCCACATCAGGCTATCCCTGT 180
 Db |||||||
 267 GTGACCCAGACCATTCATGAATATAGTGAATCCACATCCACATCAGGCTATCCCTGT 326

QY 181 GAGGAATATGAGTCCCAACTGGAATGATGGTATATCTTTCTGTTAAAGAGGATTCCTCGA 240
 Db |||||||
 327 GAGGAATATGAGTCCCAACTGGAATGATGGTATATCTTTCTGTTAAAGAGGATTCCTCGA 386

QY 241 GGCTAGTGCACCTAAGAAAGACAGGTTCCAGGCTGTGGTGTACTGAGCATGGCCTA 300
 Db |||||||
 387 GGCTAGTGCACCTAAGAAAGACAGGTTCCAGGCTGTGGTGTACTGAGCATGGCCTA 446

QY 301 GTTGGAGGTGCTAGCAACTGGATTTCACACCTGCCCAACAATAGCTGGGCTTCATTCTG 360
 Db |||||||
 447 GTTGGAGGTGCTAGCAACTGGATTTCACACCTGCCCAACAATAGCTGGGCTTCATTCTG 506

QY 361 GCAGATGCTGGTGTTCACGTGTGGATGGGAAACAGCAGGGGAAACGCTGTGTCGAAA 420
 Db |||||||
 507 GCAGATGCTGGTGTTCACGTGTGGATGGGAAACAGCAGGGGAAACGCTGTGTCGAAA 566

QY 421 CACAAGACACTCTCCATAGACCAAGATGATGCTGGGCTTCAGTTATGATGAGATGGCT 480
 Db |||||||
 567 CACAAGACACTCTCCATAGACCAAGATGATGCTGGGCTTCAGTTATGATGAGATGGCT 626

QY 481 AGGTTTGACCTTCCTGCGATGATAAACTTTATTTGACAGAAACGGGCGAGAAAGATC 540
 Db |||||||
 627 AGGTTTGACCTTCCTGCGATGATAAACTTTATTTGACAGAAACGGGCGAGAAAGATC 686

QY 541 TATTATGTGGCTATTTCAGGGCCACCATGGGCTTTTATGCAATTTCCACCATGCCA 600
 Db |||||||
 687 TATTATGTGGCTATTTCAGGGCCACCATGGGCTTTTATGCAATTTCCACCATGCCA 746

QY 601 GAGCTGGCTCAGAAAATCAAAATGATTTGCTTTAGCACCCCATGACCACTGTAAAGCAT 660
 Db |||||||
 747 GAGCTGGCTCAGAAAATCAAAATGATTTGCTTTAGCACCCCATGACCACTGTAAAGCAT 806

QY 661 GCAAAAGCCCGGGACCAAAATTTTGTGTCGCCAGATATGATGATCAAGGATGTTGTT 720
 Db |||||||
 807 GCAAAAGCCCGGGACCAAAATTTTGTGTCGCCAGATATGATGATCAAGGATGTTGTT 866

QY 721 GGCAGAAAAGAAATTTCTGTATCAGACCAATTTCTCAGACAACTTCTTATTTACCTTCT 780
 Db |||||||
 867 GGCAGAAAAGAAATTTCTGTATCAGACCAATTTCTCAGACAACTTCTTATTTACCTTCT 926

QY 781 GGCCAGGCTGATCTCTTGATCAGATTTGTAGTAATATCATCTTCTTGGTGGATTCAC 840

Db 927 GGCAGGATGCTTCTGATCAGATTTCTAGTAATATCATGTTACTTCTGGGTGATTCAC 986
 QY 841 ACCAAATATGAATGAGCCGAGCAAGTGTATGTCGCCACACACTTCTGCGAACA 900
 Db 987 ACCAAATATGAATGAGCCGAGCAAGTGTATGTCGCCACACACTTCTGCGAACA 1046
 QY 901 TCTGTGCAAAATTTCTACCTGAGCCGAGCAAGTGTATGTCGCCACACACTTCTGCGAACA 960
 Db 1047 TCTGTGCAAAATTTCTACCTGAGCCGAGCAAGTGTATGTCGCCACACACTTCTGCGAACA 1106
 QY 961 GACTGGGGAGTGAACCAAAATCTGGAATAATGCAATCAGCCAACTCTCTGTAAGGTAC 1020
 Db 1107 GACTGGGGAGTGAACCAAAATCTGGAATAATGCAATCAGCCAACTCTCTGTAAGGTAC 1166
 QY 1021 AGAGTCAGATATGACGGTCCCTACAGCAATGTGACAGGAGTGAAGTGGCTTTCA 1080
 Db 1167 AGAGTCAGATATGACGGTCCCTACAGCAATGTGACAGGAGTGAAGTGGCTTTCA 1226
 QY 1081 AATCCAGACAGCTGAAATGCTCTCTGAGGTGACCAACTCTATACCATAGAAAT 1140
 Db 1227 AATCCAGACAGCTGAAATGCTCTCTGAGGTGACCAACTCTATACCATAGAAAT 1286
 QY 1141 ATTCTGAAATGGGCTCAGTGGATTTTCATCTGGGTTTGGATCTCTCACCCTATGATAC 1200
 Db 1287 ATTCTGAAATGGGCTCAGTGGATTTTCATCTGGGTTTGGATCTCTCACCCTATGATAC 1346
 QY 1201 AATGAAATCATCTCTGATGCGAGGAGGAGCAACCTTTCCAGGAGCGTGTGAG 1260
 Db 1347 AATGAAATCATCTCTGATGCGAGGAGGAGCAACCTTTCCAGGAGCGTGTGAG 1406
 QY 1261 GCCGTATTG 1269
 Db 1407 GCCGTATTG 1415

RESULT 5
 AAD19220
 ID AAD19220 standard; DNA; 1384 BP.
 AC AAD19220;
 XX 18-DEC-2001 (first entry)
 DT Human CG162 (or C59) lipase DNA #1.
 DE Human; apolipoprotein; lipase; lipid metabolism; lipoprotein receptor; ALR; angina;
 KW cardiovascular disease; lipid metabolism; myocardial infarction;
 KW cerebral ischaemia; arterial thrombosis; thrombolytic; antilipemic;
 KW coronary artery thrombosis; cerebral artery thrombosis; stroke;
 KW intracardiac thrombosis; gene therapy; cardiovascular; vasodilator;
 KW neuroprotectant; cerebroprotective; ds.
 XX Homo sapiens.
 OS WO200179446-A2.
 XX 25-OCT-2001.
 PD 16-APR-2001; 2001WO-US12529.
 PF 14-APR-2000; 2000US-197137P.
 XX 20-JUN-2000; 2000US-0598042.
 PR 03-AUG-2000; 2000US-0631451.
 PR 22-SEP-2000; 2000US-0667298.
 PR 17-NOV-2000; 2000US-0714936.
 XX (HYSE-) HYSEQ INC.
 XX Ballinger DG, Loeb D, Montgomery JR, Tang TY, Zhou P, Goodrich R;
 PI Liu C, Asundi V, Zhao QA, Wehrman T, Drmanac RT, Ren F, Qian XB;
 PI Wang D;

DR WPI; 2001-611724/70.

XX Nucleic acids encoding human apolipoproteins, lipases, and lipoprotein

PT receptor polypeptides, useful for preventing diagnosing and treating

PT lipid metabolism disorders, thrombosis and cardiovascular diseases -

XX Claim 1; Page 160-161; 266pp; English.

PS The invention relates to polynucleotides encoding proteins CGI122, CGI179,

XX CG95, CGI21, CGI62, CG27, CGI53 and CGI168 which are related to proteins

CC involved in lipid metabolism and cardiovascular disease such as human

CC apolipoproteins, lipases and lipoprotein receptor proteins. These DNA

CC and protein sequences are useful for treating or preventing disorders

CC associated with apolipoproteins, lipases and lipoprotein receptor (ALR)

CC expression and for treating lipid metabolism, cardiovascular diseases

CC and thrombosis. Antibodies against these proteins are useful for

CC determining the presence of or predisposition to a disease associated

CC with altered levels of these sequences. ALR polypeptides are also

CC useful for identifying agents (agonists and antagonists) that bind to

CC them and cells expressing ALR proteins are useful for identifying a

CC therapeutic agent for use in treatment of a pathology related to

CC aberrant expression or physiological interactions of this polypeptide.

CC Vectors comprising these DNA and protein sequences are also useful for

CC producing ALR proteins. The nucleic acids and polypeptides of the

CC invention are also useful for the treatment of occlusive cardiovascular

CC diseases, myocardial infarction, cerebral ischaemia, angina, arterial

CC thrombosis, coronary artery thrombosis and cerebral artery thrombosis

CC or intracardiac thrombosis and stroke. The nucleotides of the invention

CC are used in gene therapy. The present sequence is human CGI162 (or C59)

CC lipase DNA.

XX SQ Sequence 1384 BP; 404 A; 281 C; 328 G; 371 T; 0 other;

Query Match 97.4%; Score 1235.4; DB 22; Length 1384;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1236; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGTTGGAACCTTGTCTGAGACAGTGGATTTCTCTCACAGAAATGGAATGCGCTCTG 60

Db 147 ATGTTGGAACCTTGTCTGAGACAGTGGATTTCTCTCACAGAAATGGAATGCGCTCTG 206

QY 61 ATTCTGTGGGGGTATATGTTCCAGAGAAATGTAATTCAGTACATATATGCAATATGCT 120

Db 207 ATTCTGTGGGGGTATATGTTCCAGAGAAATGTAATTCAGTACATATATGCAATATGCT 266

QY 121 GTGGACCCAGAGCAATTCATGAATATAGTGAATATCCTCAACATCAAGGTATCCCTGT 180

Db 267 GTGGACCCAGAGCAATTCATGAATATAGTGAATATCCTCAACATCAAGGTATCCCTGT 326

QY 181 GAGGAATATGAATGCGCAACTGAGATGGGTATATCTCTTCTTAACAGGATTCCTCGA 240

Db 327 GAGGAATATGAATGCGCAACTGAGATGGGTATATCTCTTCTTAACAGGATTCCTCGA 386

QY 241 GGCCTAGTGCACCTTAAGAGACAGGTTCCAGGCTGTGGTGTACTGACATGGCCTA 300

Db 387 GGCCTAGTGCACCTTAAGAGACAGGTTCCAGGCTGTGGTGTACTGACATGGCCTA 446

QY 301 GTTGGAGTGTAGCAACTGGATTTCCAACTGCCCAACATAGCTGGGCTTCACTCTG 360

Db 447 GTTGGAGTGTAGCAACTGGATTTCCAACTGCCCAACATAGCTGGGCTTCACTCTG 506

QY 361 GCAGATGCTGGTTTTCAGCTGTGGATGGGAAACAGCAGGGGAAACGCTGTGTCGAAA 420

Db 507 GCAGATGCTGGTTTTCAGCTGTGGATGGGAAACAGCAGGGGAAACGCTGTGTCGAAA 566

QY 421 CACAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTCAGTTATGATGAGTGGCT 480

Db 567 CACAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTCAGTTATGATGAGTGGCT 626

QY 481 AGTTTGTGCTTCTCTGAGTGAATAAATTTATTTTGCAGAAAACGGGCCAGGAAAAGATC 540

Db 627 AGTTTGTGCTTCTCTGAGTGAATAAATTTATTTTGCAGAAAACGGGCCAGGAAAAGATC 686

541 TATTATGCGCTATTCACAGGCGCACCCACCATGGCTTTTATTGCAATTTTCCACCATGCCA 600
 587 TATTATGCGCTATTCACAGGCGCACCCACCATGGCTTTTATTGCAATTTTCCACCATGCCA 746
 601 GAGCTGGCTCAGAAATCAAAATGTAATTTTGTCTTTAGCACCACCATAGCCATGTTAAGCAT 660
 747 GAGCTGGCTCAGAAATCAAAATGTAATTTTGTCTTTAGCACCACCATAGCCATGTTAAGCAT 806
 561 GCATAAGCCCGGCGCAAAATTTTGTCTGCTGCCAGATATGATCATCAAGGATTTGTT 720
 807 GCATAAGCCCGGCGCAAAATTTTGTCTGCTGCCAGATATGATCATCAAGGATTTGTT 866
 721 GGCATAAGAAATTTCTGTATCATCAGACAGATTTTCTCAGACAACTTTGTTATTACCTTTGT 780
 867 GGCATAAGAAATTTCTGTATCATCAGACAGATTTTCTCAGACAACTTTGTTATTACCTTTGT 926
 781 GGCAGGTGATTTCTGTATCATCAGATTTGTAATATCATCATGTTACTTCTGGTGATTTCAAC 840
 927 GGCAGGTGATTTCTGTATCATCAGATTTGTAATATCATCATGTTACTTCTGGTGATTTCAAC 986
 841 ACCAAATATGAATGAGCCGAGCAAGTGTATATGCTGCCCCACACACTCTTGTCTGGAACA 900
 987 ACCAAATATGAATGAGCCGAGCAAGTGTATATGCTGCCCCACACACTCTTGTCTGGAACA 1046
 901 TCTGTGAAATATTTCTACTGAGCCAGGAGTGAATTTCTGGTGAATCTCCGGCATTT 960
 1047 TCTGTGAAATATTTCTACTGAGCCAGGAGTGAATTTCTGGTGAATCTCCGGCATTT 1106
 961 GACTGGGGAGTGAGACCAAAATCTGAAATATGCAATCAGCCAACTCTCTGTAAGTAC 1020
 1107 GACTGGGGAGTGAGACCAAAATCTGAAATATGCAATCAGCCAACTCTCTGTAAGTAC 1166
 1021 AGAGTCAGATATGACGGTCCCTACAGCAATGTGAGCAGAGGTGAGGCTGGCTTTCA 1080
 1167 AGAGTCAGATATGACGGTCCCTACAGCAATGTGAGCAGAGGTGAGGCTGGCTTTCA 1226
 1081 AATCCAGAGACGTGAAATGCTGCTCTGAGGTGACCACTCATCTACCAATGAAT 1140
 1227 AATCCAGAGACGTGAAATGCTGCTCTGAGGTGACCACTCATCTACCAATGAAT 1286
 1141 ATTCTGAATGGGCTCAGTGGATTTTCACTGGGGTTTGGATGCTCTCTCAACGATGATAC 1200
 1287 ATTCTGAATGGGCTCAGTGGATTTTCACTGGGGTTTGGATGCTCTCTCAACGATGATAC 1346
 1201 AATGAATATCCATCTGATGAGCAGGAGGAGGACCA 1237
 1347 AATGAATATCCATCTGATGAGCAGGAGGAGGACCA 1383

RESULT 6
 ABS55356
 ID ABS55356 standard; cDNA; 1206 BP.
 XX AC ABS55356;
 XX DT 07-JAN-2003 (first entry)
 XX DE cDNA encoding human lipase protein.
 XX KW Human; lipase; cancer; colon; kidney; skin; brain; testis; chromosome 10;
 XX KW Burkitt's lymphoma; cytostatic; gene therapy; lipase-inhibitor; gene; ss.
 XX OS Homo sapiens.
 FH Key Location/Qualifiers
 FT 1..3 /tag= a
 FT 4..1200 /tag= b
 FT CDS /product= "Human lipase protein"
 FT 1201..1206
 FT 3'UTR /tag= c
 FT 1201..1206

PN WO200274975-A2.
 XX 26-SEP-2002.
 XX 18-MAR-2002; 2002WO-US08035.
 XX 20-MAR-2001; 2001US-0811825.
 XX (PEKE) PE CORP NY.
 PA Yan C, Di Francesco V, Beasley EM;
 PI WPI; 2002-750560/81.
 XX P-PSDB; ABG71000.
 DR New isolated human lipase peptides and encoding nucleic acids, useful
 XX for diagnosing and treating disorders mediated by human lipase
 PT proteins, e.g. cancer of the colon, kidneys, skin and testis, and
 PT Burkitt's lymphoma -
 XX Claim 4; Fig 1; 79pp; English.
 PS The present invention relates to a new lipase peptide. The methods and
 CC compositions of the present invention are useful for diagnosing and
 CC treating disorders mediated by the human lipase protein, such as cancer
 CC of the colon, kidneys, skin, brain and testis, and Burkitt's lymphoma.
 CC The present nucleic acid sequence represents the human lipase gene
 CC located on chromosome 10. This sequence encodes the human lipase protein
 CC of the invention.
 XX Sequence 1206 BP; 339 A; 260 C; 295 G; 312 T; 0 other;
 SQ Query Match 88.4%; Score 1122; DB 24; Length 1206;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 148 AGTGAATATCATCAACATCAAGGCTATCCCTGTGAGGAATATGAGTCGCAACTGAAGAT 207
 DB 76 AGTGAATATCATCAACATCAAGGCTATCCCTGTGAGGAATATGAGTCGCAACTGAAGAT 135
 QY 208 GGGTATATCTTTCTGTTAAACAGGATTCCTCGAGCCCTAGTGCACCTAAGACAGGAT 267
 DB 136 GGGTATATCTTTCTGTTAAACAGGATTCCTCGAGCCCTAGTGCACCTAAGACAGGAT 195
 QY 268 TCCAGGCTGTGTTTACTGAGCATGCGCTAGTTGAGGTTGCTAGCACTGGATTTC 327
 DB 196 TCCAGGCTGTGTTTACTGAGCATGCGCTAGTTGAGGTTGCTAGCACTGGATTTC 255
 QY 328 AACCTGCCCAACAATAGCTGGCTTCATTTCTGGCAGATGCTGTTTGAAGTGGATG 387
 DB 256 AACCTGCCCAACAATAGCTGGCTTCATTTCTGGCAGATGCTGTTTGAAGTGGATG 315
 QY 388 GGGACAGCAGGGGAAACGCTGTTCTGAAACACAGACACTCTCCATAGACCAAGAT 447
 DB 316 GGGACAGCAGGGGAAACGCTGTTCTGAAACACAGACACTCTCCATAGACCAAGAT 375
 QY 448 GAGTCTGGGCTTTTTCAGTTATGATGAGATGGCTAGGTTTGAAGTGGATGATAAC 507
 DB 376 GAGTCTGGGCTTTTTCAGTTATGATGAGATGGCTAGGTTTGAAGTGGATGATAAC 435
 QY 508 TTTATTTTGCAGAAACGGGCGCAGGAAGATCTATTATGTCGCTATTTCAGGGCACC 567
 DB 436 TTTATTTTGCAGAAACGGGCGCAGGAAGATCTATTATGTCGCTATTTCAGGGCACC 495
 QY 568 ACCATGGGCTTTTATGCTATTTTCCACCATGCCAGAGCTGGCTCAGAAATCAAAATGAT 627
 DB 496 ACCATGGGCTTTTATGCTATTTTCCACCATGCCAGAGCTGGCTCAGAAATCAAAATGAT 555
 QY 628 TTTGCTTTAGCACCACCATAGCCATGTTAAGCATGCAAAAAGCCCGGACCAATTTT 687
 DB 556 TTTGCTTTAGCACCACCATAGCCATGTTAAGCATGCAAAAAGCCCGGACCAATTTT 615
 QY 688 TTGCTGCCAGATATGATCAAGGATTTGTTGGCAAAAAGAAATTTCTGTATCAGACC 747

493	ACATGGCGTTTATGTGCAATTTTCCACATGCCAGAGCTGGCTCAGAAAAATCAAAATGTAT	552
628	TTTGTCTTTAGCACCCCATAGCCACTGTTTAAGCATGCAAAAAGCCCGGGACCAAAATTTTG	687
553	TTTGTCTTTAGCACCCCATAGCCACTGTTTAAGCATGCAAAAAGCCCGGGACCAAAATTTTG	612
688	TTGCTGCCAGATATGATCAAGAGGATTTGTTTGGCNAAAAAGAAATTCCTGTATCAGACC	747
613	TTGCTGCCAGATATGATCAAGAGGATTTGTTTGGCNAAAAAGAAATTCCTGTATCAGACC	672
748	AGATTTCTCAGACAACTGTGTATTACTTTGTGGCCAGGTGATCTCTTGATCAGATTTGT	807
673	AGATTTCTCAGACAACTGTGTATTACTTTGTGGCCAGGTGATCTCTTGATCAGATTTGT	732
808	AGTAATATCATGTTACTCTGGGTGGATCAACACCAACAATATGAACATGAGCCGAGCA	867
733	AGTAATATCATGTTACTCTGGGTGGATCAACACCAACAATATGAACATGAGCCGAGCA	792
868	AGTGTATATCTGCCACACCTCTGTCTGGAACATCTGTGCAAAATATTCCTACACTGGAGC	927
793	AGTGTATATCTGCCACACCTCTGTCTGGAACATCTGTGCAAAATATTCCTACACTGGAGC	852
928	CAGGCAGTGAATTTCTGTGTGAACCTCCGGCCATTTGAATCTGGGGGAGTGAGACCAAAAATCTG	987
853	CAGGCAGTGAATTTCTGTGTGAACCTCCGGCCATTTGAATCTGGGGGAGTGAGACCAAAAATCTG	912
988	GAATAATGCATCAGCCAACTCTCTGTTAAGGTACAGAGTCAGAGATATGACGGTCCCTACA	1047
913	GAATAATGCATCAGCCAACTCTCTGTTAAGGTACAGAGTCAGAGATATGACGGTCCCTACA	972
1048	GCAATGTGGACAGGAGGTCAGGACTGGCTTTCAAATCCAGAAGACGCTGAAAATGCTGCTC	1107
973	GCAATGTGGACAGGAGGTCAGGACTGGCTTTCAAATCCAGAAGACGCTGAAAATGCTGCTC	1032
1108	TCTGAGGTGACCAACCTCATCTACCAATAGAAATATTCCTGAATGGGCTCAGCTGGATTTTC	1167
1033	TCTGAGGTGACCAACCTCATCTACCAATAGAAATATTCCTGAATGGGCTCAGCTGGATTTTC	1092
1168	ATCTGGGGTTTGGATGCTCTCACCGTATGTACAAATGAAATCATCCATCTGTATGCACGAG	1227
1093	ATCTGGGGTTTGGATGCTCTCACCGTATGTACAAATGAAATCATCCATCTGTATGCACGAG	1152
1228	GAGGAGACCAACCTTTTCCCAGGGACGGTGTGAGGCCGTATTG	1269
1153	GAGGAGACCAACCTTTTCCCAGGGACGGTGTGAGGCCGTATTG	1194

DEC 11 1964

RESUL 8
ABK90871

ABK90871
ID ABK90871 standard: cDNA: 1272 BP

AC
BX
BX90871.

XX
DT 05-NOV-2002 (first entry)

XX
the cDNA encoding human lipid-associated molecule 1 (LAPM1)

XX Human; lipid-associated molecule; LIPAM; cancer; leukaemia; gene; ss;
 KW cardiovascular disorder; atherosclerosis; hypertension; stroke;
 KW angina pectoris; neurological disorder; Alzheimer's disease; epilepsy;
 KW dementia; Parkinson's disease; multiple sclerosis; anxiety; psoriasis;
 KW autoimmune disorder; inflammatory disorder; anaemia; asthma; trauma;
 KW acquired immunodeficiency syndrome; AIDS; bronchitis; diabetes mellitus;
 KW emphysema; Crohn's disease; atopic dermatitis; rheumatoid arthritis;
 KW gastrointestinal disorder; lipid metabolism disorder; transgenic animal.

Homo sapiens.

XX	Kay
FH	

FT	KEY	CDS
----	-----	-----

CD3
CD4
CD8
CD11b
CD11c
CD13
CD14
CD15
CD16
CD18
CD19
CD20
CD22
CD23
CD24
CD25
CD26
CD27
CD28
CD29
CD30
CD31
CD32
CD33
CD34
CD35
CD36
CD37
CD38
CD39
CD40
CD41
CD42
CD43
CD44
CD45
CD46
CD47
CD48
CD49
CD50
CD51
CD52
CD53
CD54
CD55
CD56
CD57
CD58
CD59
CD60
CD61
CD62
CD63
CD64
CD65
CD66
CD67
CD68
CD69
CD70
CD71
CD72
CD73
CD74
CD75
CD76
CD77
CD78
CD79
CD80
CD81
CD82
CD83
CD84
CD85
CD86
CD87
CD88
CD89
CD90
CD91
CD92
CD93
CD94
CD95
CD96
CD97
CD98
CD99
CD100
CD101
CD102
CD103
CD104
CD105
CD106
CD107
CD108
CD109
CD110
CD111
CD112
CD113
CD114
CD115
CD116
CD117
CD118
CD119
CD120
CD121
CD122
CD123
CD124
CD125
CD126
CD127
CD128
CD129
CD130
CD131
CD132
CD133
CD134
CD135
CD136
CD137
CD138
CD139
CD140
CD141
CD142
CD143
CD144
CD145
CD146
CD147
CD148
CD149
CD150
CD151
CD152
CD153
CD154
CD155
CD156
CD157
CD158
CD159
CD160
CD161
CD162
CD163
CD164
CD165
CD166
CD167
CD168
CD169
CD170
CD171
CD172
CD173
CD174
CD175
CD176
CD177
CD178
CD179
CD180
CD181
CD182
CD183
CD184
CD185
CD186
CD187
CD188
CD189
CD190
CD191
CD192
CD193
CD194
CD195
CD196
CD197
CD198
CD199
CD200
CD201
CD202
CD203
CD204
CD205
CD206
CD207
CD208
CD209
CD210
CD211
CD212
CD213
CD214
CD215
CD216
CD217
CD218
CD219
CD220
CD221
CD222
CD223
CD224
CD225
CD226
CD227
CD228
CD229
CD230
CD231
CD232
CD233
CD234
CD235
CD236
CD237
CD238
CD239
CD240
CD241
CD242
CD243
CD244
CD245
CD246
CD247
CD248
CD249
CD250
CD251
CD252
CD253
CD254
CD255
CD256
CD257
CD258
CD259
CD260
CD261
CD262
CD263
CD264
CD265
CD266
CD267
CD268
CD269
CD270
CD271
CD272
CD273
CD274
CD275
CD276
CD277
CD278
CD279
CD280
CD281
CD282
CD283
CD284
CD285
CD286
CD287
CD288
CD289
CD290
CD291
CD292
CD293
CD294
CD295
CD296
CD297
CD298
CD299
CD300
CD301
CD302
CD303
CD304
CD305
CD306
CD307
CD308
CD309
CD310
CD311
CD312
CD313
CD314
CD315
CD316
CD317
CD318
CD319
CD320
CD321
CD322
CD323
CD324
CD325
CD326
CD327
CD328
CD329
CD330
CD331
CD332
CD333
CD334
CD335
CD336
CD337
CD338
CD339
CD340
CD341
CD342
CD343
CD344
CD345
CD346
CD347
CD348
CD349
CD350
CD351
CD352
CD353
CD354
CD355
CD356
CD357
CD358
CD359
CD360
CD361
CD362
CD363
CD364
CD365
CD366
CD367
CD368
CD369
CD370
CD371
CD372
CD373
CD374
CD375
CD376
CD377
CD378
CD379
CD380
CD381
CD382
CD383
CD384
CD385
CD386
CD387
CD388
CD389
CD390
CD391
CD392
CD393
CD394
CD395
CD396
CD397
CD398
CD399
CD400
CD401
CD402
CD403
CD404
CD405
CD406
CD407
CD408
CD409
CD410
CD411
CD412
CD413
CD414
CD415
CD416
CD417
CD418
CD419
CD420
CD421
CD422
CD423
CD424
CD425
CD426
CD427
CD428
CD429
CD430
CD431
CD432
CD433
CD434
CD435
CD436
CD437
CD438
CD439
CD440
CD441
CD442
CD443
CD444
CD445
CD446
CD447
CD448
CD449
CD450
CD451
CD452
CD453
CD454
CD455
CD456
CD457
CD458
CD459
CD460
CD461
CD462
CD463
CD464
CD465
CD466
CD467
CD468
CD469
CD470
CD471
CD472
CD473
CD474
CD475
CD476
CD477
CD478
CD479
CD480
CD481
CD482
CD483
CD484
CD485
CD486
CD487
CD488
CD489
CD490
CD491
CD492
CD493
CD494
CD495
CD496
CD497
CD498
CD499
CD500
CD501
CD502
CD503
CD504
CD505
CD506
CD507
CD508
CD509
CD510
CD511
CD512
CD513
CD514
CD515
CD516
CD517
CD518
CD519
CD520
CD521
CD522
CD523
CD524
CD525
CD526
CD527
CD528
CD529
CD530
CD531
CD532
CD533
CD534
CD535
CD536
CD537
CD538
CD539
CD540
CD541
CD542
CD543
CD544
CD545
CD546
CD547
CD548
CD549
CD550
CD551
CD552
CD553
CD554
CD555
CD556
CD557
CD558
CD559
CD560
CD561
CD562
CD563
CD564
CD565
CD566
CD567
CD568
CD569
CD570
CD571
CD572
CD573
CD574
CD575
CD576
CD577
CD578
CD579
CD580
CD581
CD582
CD583
CD584
CD585
CD586
CD587
CD588
CD589
CD590
CD591
CD592
CD593
CD594
CD595
CD596
CD597
CD598
CD599
CD600
CD601
CD602
CD603
CD604
CD605
CD606
CD607
CD608
CD609
CD610
CD611
CD612
CD613
CD614
CD615
CD616
CD617
CD618
CD619
CD620
CD621
CD622
CD623
CD624
CD625
CD626
CD627
CD628
CD629
CD630
CD631
CD632
CD633
CD634
CD635
CD636
CD637
CD638
CD639
CD640
CD641
CD642
CD643
CD644
CD645
CD646
CD647
CD648
CD649
CD650
CD651
CD652
CD653
CD654
CD655
CD656
CD657
CD658
CD659
CD660
CD661
CD662
CD663
CD664
CD665
CD666
CD667
CD668
CD669
CD670
CD671
CD672
CD673
CD674
CD675
CD676
CD677
CD678
CD679
CD680
CD681
CD682
CD683
CD684
CD685
CD686
CD687
CD688
CD689
CD690
CD691
CD692
CD693
CD694
CD695
CD696
CD697
CD698
CD699
CD700
CD701
CD702
CD703
CD704
CD705
CD706
CD70

$\Gamma_{\text{product}}/\Gamma_{\text{cag}} =$

1
2
3
3
3
0
4
4

W0300246478-A2

13-JUN-2002

04-DEC-2001: 2001WQ-IIS47430-

08-DEC-2000: 2000US-254505P.

08-DEC-2000; 2000US-254505P.
15-DEC-2000; 2000US-256187P.

13-DEC-2000; 2000US-236167F.
22-DEC-2000: 2000US-257908P.

22-DEC-2000; 2000US-237308F
26-JAN-2001; 2001US-264429P.

(TNCY-) INCYTE GENOMICS INC.

Griffin, J.A.	Gandhi, A.P.	Ramkumar, J.	Tang, Y.T.	Ding, L.	Yue, H.
---------------	--------------	--------------	------------	----------	---------

GILLIN JA, GANQUH AR, KAMKUMAL U, TANG IL, BRING H, RUE H, SAPPERSTEIN SK, HONCHELL CD, BRUNS CM, DUGGAN BM; Gietzen KJ.

Xu Y., Lee S:

WPT: 2002-619022/66

WPI; Z00Z-6I30ZZ/
P=PSNB: ABG3189A

Novel polypeptide of human lipid associated molecule, useful for diagnosing, treating and preventing cancer, hypercholesterolemia, cirrhosis, myocardial infarction, Parkinson's disease, asthma, psoriasis. gastritis -

Claim 5: Page 120-121: 126pp: English:

The invention relates to an isolated polypeptide (I) of human lipid-associated molecule (LIPAM), and an isolated polynucleotide (II) encoding (I). (I), (II) and antibody to (I) are useful for treating a disease or condition associated with decreased or increased expression of functional LIPAM. (I) or (II) is useful for diagnosing, treating or preventing cancer e.g. leukaemia, cancer of adrenal gland, bladder, bone, bone marrow, brain, ovary, etc.; cardiovascular disorders e.g. atherosclerosis, hypertension, Raynaud's disease, angina pectoris, myocardial infarction, rheumatic fever, cardiomyopathy, pericarditis, chronic obstructive pulmonary disease, etc; neurological disorders e.g. Alzheimer's disease, stroke, epilepsy, dementia, Parkinson's disease, multiple sclerosis, Creutzfeldt-Jakob disease, myasthenia gravis, anxiety, diabetic neuropathy, etc; autoimmune/inflammatory disorders e.g. acquired immunodeficiency syndrome (AIDS), Addison's disease, anaemia, asthma, bronchitis, diabetes mellitus, emphysema, Crohn's disease, atopic dermatitis, rheumatoid arthritis, psoriasis, systemic lupus erythematosus, trauma, etc; and gastrointestinal disorders e.g. gastritis, anorexia, nausea gastroenteritis, ulcerative colitis, cholecystitis, hepatitis, cirrhosis, hepatoma, diarrhoea, constipation, etc, and disorders of lipid metabolism e.g., fatty liver, cholestasis, Fabry's disease, hypercholesterolaemia. (I) or (II) is useful for assessing the effects of exogenous compounds on the expression of (I) or (II). (I) is useful in number of drug screening techniques, and to analyse the proteome of a tissue or cell type. (I) or antibody to (II) is useful as elements on a microarray. (II) is useful for creating knockin humanized animals or transgenic animals to model human diseases, in somatic or germline gene therapy, to generate a transcript image of a tissue or cell type, for detecting differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier or affected individuals, and as hybridisation probes for mapping naturally occurring genomic sequences. The present sequence represents the coding sequence of a human lipid-associated molecule (LIPAM) of the invention.

Semence 1272 BP: 351 A: 263 C: 311 G: 347 T: 0 other:

over Math

Very Match	87.1%	Score 1104.8;	Length 1272;
at local similarity	87.2%	Pred No 0;	

Best Local Similarity 97.2%; Pred. No. 0; Mismatches 1143; Conservative 0

115 AAAGCTGTGGACCCAGAAGCATTCATGAATATTAGTGAAATCATCCACATCAAGGCTAT 174

94 AATGAACTGAATCCTGACGCTGCGATCGAATACCTAGTGAAATCATCCACATCAAGCGTAT 153

175 C C T C T C C C A A T A T C T C C C A A C C G G A T T 234

I/S CCCGTCGAGGAAATATGAAGTCGCAACTGAAGATGGGTAATCCCTTCGTAAACAGGAT 234

235 CCTCGAGGCTAGTGCACCACTAAGACACAGGTTCCAGGCTGTGCTTACTGACGAT 294
 214 CTTGAGGCTAGTGCACCACTAAGACACAGGTTCCAGGCTGTGCTTACTGACGAT 273
 295 GGCCTAGTTGGAGGTGCTAGCACTGGAATTTCCAACTGCCCCCAATAGCTGGCTTC 354
 274 GGCCTAGTTGGAGGTGCTAGCACTGGAATTTCCAACTGCCCCCAATAGCTGGCTTC 333
 355 ATCTGGCAGAGTCTGTTTACCTGTGATGGGGAACAGCAGGGGAAACGCTGCT 414
 334 ATCTGGCAGAGTCTGTTTACCTGTGATGGGGAACAGCAGGGGAAACGCTGCT 393
 415 CGAAAAACAAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTTCAGTTATGATGAG 474
 394 CGAAAAACAAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTTCAGTTATGATGAG 453
 475 ATGGCTAGTTGACCTCTCGAGTATGAATTTTTCAGAAACGGGCGGAGAA 534
 454 ATGGCTAGTTGACCTCTCGAGTATGAATTTTTCAGAAACGGGCGGAGAA 513
 535 AAGATCTATTATGCTGGCTATTTCACAGGGCACCACCATGGCTTTTATTCATTTTCCACC 594
 514 AAGATCTATTATGCTGGCTATTTCACAGGGCACCACCATGGCTTTTATTCATTTTCCACC 573
 595 ATGCCAGAGTCTGCTCAGAAATCAAAATGATATTTTGTCTTATGACCCATAGCCACTGTT 654
 574 ATGCCAGAGTCTGCTCAGAAATCAAAATGATATTTTGTCTTATGACCCATAGCCACTGTT 633
 655 AAGCATGCAAAAGCCCGGACCAAAATTTTGTCTGCTCAGATATGATCAAGGA 714
 634 AAGCATGCAAAAGCCCGGACCAAAATTTTGTCTGCTCAGATATGATCAAGGA 693
 715 TTGTTGGCAAAAGAAATTTTGTATCAGACAGATTTCTCAGACAACTTGTTATTTAC 774
 694 TTGTTGGCAAAAGAAATTTTGTATCAGACAGATTTCTCAGACAACTTGTTATTTAC 753
 775 CTTTGTGGCAGGTGATCTTGTATCAGATTTGATTAATATCATGTTCTTGGGTGA 834
 754 CTTTGTGGCAGGTGATCTTGTATCAGATTTGATTAATATCATGTTCTTGGGTGA 813
 835 TTCAACACCAACATATGAACAT-----GAGCCGAGCAAGTGA 873
 814 TTCAACACCAACATATGAACATGATCTGTTTGTACAGAGCCGAGCAAGTGA 873
 874 TATGCTGCCACACTCTGCTGGAACATCTGTGCAAAATTTCTACTGGAGCCAGCA 933
 874 TATGCTGCCACACTCTGCTGGAACATCTGTGCAAAATTTCTACTGGAGCCAGCA 933
 934 GTGAATCTGTGAACCTCGGGCACTTTGACTGGGGAGTGAGACCAAAATCTGGAATA 993
 934 GTGAATCTGTGAACCTCGGGCACTTTGACTGGGGAGTGAGACCAAAATCTGGAATA 993
 994 TGCAATCAGCAACTCTGTTAAGTACAGAGTACAGATATGAGGTCCTTACAGCAATG 1053
 994 TGCAATCAGCAACTCTGTTAAGTACAGAGTACAGATATGAGGTCCTTACAGCAATG 1053
 1054 TGGACAGAGTCTGAGGCTTTCAATCCAGAGACGTGAATGCTGCTCTGAG 1113
 1054 TGGACAGAGTCTGAGGCTTTCAATCCAGAGACGTGAATGCTGCTCTGAG 1113
 1114 GTGACCAACCTCATCTACCATAGAAATTTCTGAATGGGCTACGTTGATTTTCATCTGG 1173
 1114 GTGACCAACCTCATCTACCATAGAAATTTCTGAATGGGCTACGTTGATTTTCATCTGG 1173
 1174 GGTGTTGATGCTCTCCAGGCTATGATCAATCAATCATCTGATGTCAGCAGGAGGAG 1233
 1174 GGTGTTGATGCTCTCCAGGCTATGATCAATCAATCATCTGATGTCAGCAGGAGGAG 1233
 1234 ACCAACCTTTCCAGGGACGGTGTGAGCCGTTATG 1269
 1234 ACCAACCTTTCCAGGGACGGTGTGAGCCGTTATG 1269

RESULT 9
 ABK86570
 ID ABK86570 standard; cDNA; 1116 BP.
 XX
 AC ABK86570;
 XX
 DT 24-SEP-2002 (first entry)
 XX
 DE Human cDNA encoding lysosomal acid lipase #2.
 XX
 KW Human; ss; gene; lysosomal acid lipase; lipid malabsorption illness;
 KW cystic fibrosis; alcoholism; heart disease; heart attack;
 KW Wolman disease; cholesterol ester storage disease; brain injury;
 KW mood disorder; anxiety disorder; thought disorder; volition disorder;
 KW sleep disorder; neurogenic disorder; myopathic disorder; COPD;
 KW obesity; cancer; chronic obstructive pulmonary disease; diabetes;
 KW cardiovascular disorder; Alzheimer's disease; Parkinson's disease;
 KW anorexia; osteoarthritis; central nervous system disorder;
 KW peripheral nervous system disorder.
 XX
 OS Homo sapiens.
 XX
 Key Location/Qualifiers
 CDS 1..1116
 FT /*tag= a
 FT /product= "Lysosomal acid lipase #2"
 XX
 PN WO200236731-A2.
 XX
 PD 10-MAY-2002.
 XX
 PF 30-OCT-2001; 2001WO-EPI2518.
 XX
 PR 31-OCT-2000; 2000US-244170P.
 PR 29-MAY-2001; 2001US-293516P.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Xiao Y;
 XX
 DR WPI; 2002-519248/55.
 DR P-PSDB; AAU99165.
 XX
 PT Novel human lysosomal acid lipase polypeptide, useful for treating
 PT cancer, diabetes, obesity, chronic obstructive pulmonary disease,
 PT peripheral or central nervous system disorder or cardiovascular
 PT disorder -
 XX
 PS Claim 1; Fig 4; 126pp; English.
 XX
 CC The invention relates to a purified human lysosomal acid lipase
 CC polypeptide. Also included are the polynucleotide encoding the
 CC lipase (or its fragment, derivative, allele or sequence at least 60%
 CC identical to it), vectors, host cells, a reagent (e.g. an antisense
 CC oligonucleotide) which binds to the lipase or polynucleotide (used for
 CC detection and modulating/reducing the lipase activity) and an anti-lipase
 CC antibody. The lipase and polynucleotide are useful for identifying
 CC therapeutic agents that either increase or decrease the lipase activity.
 CC The identified agent, the lipase and polynucleotide are useful for
 CC treatment of a disease such as lipid malabsorption illness,
 CC cystic fibrosis, alcoholism, heart disease, heart attack,
 CC Wolman disease, cholesterol ester storage disease, brain injury,
 CC mood disorder, anxiety disorder, thought disorder, volition disorder,
 CC sleep disorder, neurogenic disorder, myopathic disorder, obesity, cancer,
 CC chronic obstructive pulmonary disease (COPD), diabetes,
 CC cardiovascular disorder, Alzheimer's disease, Parkinson's disease,
 CC anorexia, osteoarthritis, a central nervous system disorder and
 CC a peripheral nervous system disorder. The present sequence is the
 CC cDNA encoding human lysosomal lipase #2.
 XX
 SQ Sequence 1116 BP; 326 A; 234 C; 245 G; 311 T; 0 other;

472 ATGGCAAAATATGACCTACAGCTTCCATTAACCTTCAATTAATAAACTGGCCCAAGAA 531
 535 AAGATCTATTATGCTGGCTATTACAGGGCACCACCATGGCTTTATTCATTTCCACC 594
 532 CAAGTGATATGTTGGTCAATTCAGAGGCACCATATAGTTTATAGCATTTTCACAG 591
 595 ATGCCAGAGCTGGCTCAGAAAATCAAAATATATTTTGGCTTTAGCACCACCATAGCTGTT 654
 592 ATCCCTGAGCTGGCTAAAGGATTAATAATGTTTTTGGCCCTGGCTCTGGCTTCCGTC 651
 655 AAGCATCAAAAGCCCGGACCAATTTTGTGCTGCCAGATATGATCAAGGA 714
 652 GCCTTCTGTACTAGCCCTATGGCCAAATTAGGACGATTAACAGATCATCTCATTAAGGAC 711
 715 TTGTTTGGCAAAAGAAATTTCTGTATCAGACCAGATTTCTCAGACAACTTGTTATTAC 774
 712 TTATTTGGACAAAGATTTCTTCCAGAGTGGTTTTTGAAGTGGCTGGGTACCCAC 771
 775 CTTTGTGGCAGGATGATCTGTATCAGATTTGTAGTAAATATCATGTTACTTCTGGGTGA 834
 772 GTTTGCACTCATGTACTAGAGGAGCTCTGTGGAATCTCTGTTTTCTTCTGTGTGA 831
 835 TTCAACACCAACATATGAACATGAGCCGAGCAAGTGTATATCTGCCACACTCTTGCT 894
 832 TTTATGAGAGAAATTTAAATATGCTAGAGTGGATGATATACACATTTCTCTGCT 891
 895 GGAACATCTGTGCAAAATTTCTACACTGGAGCCAGCAGTGAATTTCTGGTGAATCCCG 954
 892 GGAACATCTGTGCAAAATTTCTACACTGGAGCCAGCCTGTTAAATTTCCAAAAGTTTCAA 951
 955 GCATTTGACCTGGGAGTGAGACCAAAATCTGMAAAATGCAATCAGCCACTCTGTA 1014
 952 GCCTTTGACCTGGGAGTGAGACCAAAATTTCTACACTGGAGCCAGCAGTGAATTTCTGGTGAATCCCG 1011
 1015 AGGTACAGAGTACAGATATGACCGTCTCTACAGCAATGTGGACAGAGTCAAGACTG 1074
 1012 ACATACATGTGAAGGACATGCTGTGCCGACTGCACTCTGGAGCGGGGTACGACTG 1071
 1075 CTTTCAATCCAGAGAGCGTGAATATGCTGCTCTGAGTGAGCAACCTCATCTACCAT 1134
 1072 CTTGAGATGCTACGAGCTCAATATCTTACTGACTCAGATCAACCACTTGGTGTCCAT 1131
 1135 AAGAAATTTCTGAAATGGGCTCAGTGGATTTCTATCTGGGCTTTGGATGCTCCCTACCGT 1194
 1132 GAGAGCATTCGGAAATGGAGCATTTGACTTCTTGGGCTGGATGCGCTTGGAGG 1191
 1195 ATGTACAAATGAATCATCTTCATGTCAGAGGAGGAGCAACCT 1241
 1192 CTTTATAATAAAATTTAATCTAATGAGGAAATATCAGTGAAGCT 1238

RESULT 11
 AAH02906 standard; DNA; 2493 BP.
 ID AAH02906
 XX
 AC AAH02906;
 XX
 DT 15-JUN-2001 (first entry)
 XX
 DE Human shear stress-response coding sequence SEQ ID NO: 65.
 XX
 KW Human; shear stress-response protein; vascular disease;
 XX arteriosclerosis; ds.
 OS Homo sapiens.
 PN WO200125427-A1.
 XX
 PD 12-APR-2001.
 XX
 PF 02-OCT-2000; 2000WO-JP06840.
 XX
 PR 01-OCT-1999; 99JP-0280976.

XX (KYOW) KYOWA HAKKO KOGYO KK.
 PA (NOJII) NOJIMA H.
 XX Nojima H, Yoshisue H, Obayashi M, Ota T, Kawabata A, Sakurada K;
 PI Kuga T, Sekine S, Nakamura Y, Sugano S;
 XX WPI; 2001-266308/27.
 DR P-PSDB; AAB90783.
 XX
 PT DNA sequences, proteins encoded by them and antibodies against them
 PT are used in diagnosis and treatment of vascular disease caused by
 PT arteriosclerosis -
 XX
 PS Claim 20; Page 398-402; 678pp; Japanese.
 XX
 CC The present invention provides the protein and coding sequences of a
 CC number of human shear stress response proteins. These are useful in the
 CC diagnosis, treatment and screening of vascular diseases caused by
 CC arteriosclerosis, including heart failure, post-PTCA restenosis and
 CC hypertension.
 XX
 SQ Sequence 2493 BP; 692 A; 484 C; 516 G; 801 T; 0 other;
 Query Match 41.5%; Score 526.2; DB 22; Length 2493;
 Best Local Similarity 66.6%; Pred. No. 1.3e-156;
 Matches 751; Conservative 0; Mismatches 376; Indels 0; Gaps 0;
 QY 115 AAAGCTGTGGACCCAGAGCAATTCATGAATATTTAGTGAATATCCAAATCAAGGCTAT 174
 DB 119 ACAGCTGTGGATCTGAAACAAACATGATGTGAGTGAATATCTCTTACTGGGATTC 178
 QY 175 CCCTGTGAGGAATATGAGTGCCTGCACTGAGATGGGTATATCTCTTCTGTAAACAGATT 234
 DB 179 CCTAGTGGGAATACCTAGTTGAGACAGAGATGATATATCTGTGCTTAAACCGAAT 238
 QY 235 CCTGAGGCTTAGTGCAACCTTAAGAGACAGAGTTCAGGCTGTGGTGTACTGACAGCAT 294
 DB 239 CCTCATGGGAGGAAGAACCACTTCTGACAAAGTCCCAACAGTTGTCTTCTGCAACAT 298
 QY 295 GGCCTAGTTGAGGTGTAGCACTGATTTCCAACTGCACTGCACTGCACTGCACTGCTTC 354
 DB 299 GCTTGTCTGGCAGATTTCTAGTAACTGGTCAAACTTGGCAACAGCAGCCTGGGCTTC 358
 QY 355 ATCTGGCAGATGCTGTTTGGCTGTGGATGGGAAACAGCAGGGAACAGCCTGCTCT 414
 DB 359 ATCTGCTGATGCTGTTTGGCTGTGGATGGGAAACAGCAGGGAACAGCCTGCTCT 418
 QY 415 CGAAACACAAAGACACTCTCCATAGACCAAGATGAGTTCGGGCTTTTCAATATGATGAG 474
 DB 419 CGGAAACATAAGACACTCTCAGTCTTCTCAGATGAATTTCTGGGCTTTTCAATATGATGAG 478
 QY 475 ATGGCTAGTTTGGCTTCTGCTGAGTGAATATTTTGGCAAGAAACGGGCGCAGAA 534
 DB 479 ATGGCAAAATATGACCTTACCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 538
 QY 535 AAGATCTATTATGCTGGCTATTTCAGGGCAACCACTGCTTATTTGCAATTTTCCACC 594
 DB 539 CAAGTGTATTATGCTGGGCTATTCTCAAGGCAACCACTATAGTTTATAGCATTTTACAG 598
 QY 595 ATGCCAGAGCTGGCTCAGAAAATCAAAATGATGATTTTCTTTAGCACCACCATAGCCTGTT 654
 DB 599 ATCCCTGAGCTGGCTAAAGGATTAATAATGTTTTTGGCCCTGGCTCTGGCTTCCGTC 658
 QY 655 AAGCATCAAAAGCCCGGACCAATTTTGTGCTGCCAGATATGATCAAGGGA 714
 DB 659 GCCTTCTGTACTAGCCCTATGGCCAAATTAGGACGATTAACAGATCATCTCATTAAGGAC 718
 QY 715 TTGTTTGGCAAAAGAAATTTCTGTATCAGACAGATTTCTCAGACAACTTGTATTATAC 774
 DB 719 TTATTTGGAGCAAAAGAAATTTCTTCCCGAGAGTGGCTTTTGAAGTGGCTGGTACCCAC 778
 QY 775 CTTTGTGGCCAGGTGATCTTGTATGATGATTTGTAGTATATATATATCTTACTTCTGGTGA 834

Db 779 GTTTCACCTCATGTCTAAGAGAGCTGTGTGAAATCTCTGTTCTCTGTGTGGA 838
 Qy 835 TTCAACACCAAAATATGAACATGAGCGAGCAAGTGTATATGCTGCCACACATCTTGCT 894
 Db 839 TTTATGAGAGAAATTTAAATATGCTAGAGTGGATGATATACAAACATCTCTGCT 898
 Qy 895 GGAACATCTGTGCAAAATATTCTACATGGAGCCAGGAGTGAATTTCTGTGAACTCCGG 954
 Db 899 GGAACCTTCTGTGCAAAACATGTTACATGGAGCCAGGCTGTTAAATTCCAAAGTTTCAA 958
 Qy 955 GCATTTGACTGGGAGTGGAGACCAAAATCTGGAATAATGCAATCAGCCAACTCTGTGA 1014
 Db 959 GCGTTTGACTGGGAGAGCAGTCCCAAGAAATTTTTCATTACACAGAGTTA'TCCCTCCC 1018
 Qy 1015 AGGTACAGAGTCAGAGATATGAGCTCCCTACGCAATGTGACAGAGGTCAGGACTGG 1074
 Db 1019 ACATACAAATGTGAGGACATGCTTGTGCGAGCTGCAGTCTGGAGCGGGGTACAGACTGG 1078
 Qy 1075 CTTTCAATCCAGAGAGCTGAAATGCTGCTCTGAGGTGACCAACTCATCTACCAT 1134
 Db 1079 CTTGAGATGTCTACGAGCTCAATATCTTACTGACTCAGATCACCACCTTGGTGTCCAT 1138
 Qy 1135 AAGAATATTCCTGAATGGCTCAGCTGGATTTCACTCGGGTTTGGATGCTCTCCACCGT 1194
 Db 1139 GAGAGCATTCGGAAATGGAGCATCTTGACTTCATTTGGGGCTGGATGCCCTTGGAGG 1198
 Qy 1195 ATGTCAATGAAATCATCATCTGATGAGCAGGAGGAGACCAACT 1241
 Db 1199 CTTTATAATAAATTAATTAATCTAATGAGAAATATCAGTGAAGCT 1245

RESULT 12

AAQ42310
 ID AAQ42310 standard; DNA; 1378 BP.
 AC AAQ42310;
 XX
 DT 25-MAR-2003 (updated)
 DT 20-SEP-1993 (first entry)
 XX
 DE RGL precursor.
 XX
 KW Rabbit gastric lipase; RGL; prGLN2.1; fat; bioconversion;
 KW hydrolysis; transesterification; ds.
 XX
 OS Oryctolagus cuniculus.

XX
 FH Key Location/Qualifiers
 FT misc_RNA 120...1247
 FT /tag= a
 FT /note= "claim 1; page 10-11"
 XX
 PN EP542629-A1.
 XX
 PD 19-MAY-1993.
 XX
 PF 12-NOV-1992; 92EP-0403055.
 XX
 PR 13-NOV-1991; 91FR-0013948.
 XX
 PA (LJOU) INST RECH JOUVEINAL.
 XX
 PI Benicourt C, Blanchard C, Junien J;
 XX
 DR WPI; 1993-161080/20.
 DR P-PSDB; AAR37302.
 XX
 PT Rabbit gastric lipase, its precursor and their DNA - useful for
 PT treating conditions linked to gastric lipase deficiency, such as
 PT mucoviscidiosis and pancreatic exocrine insufficiency
 XX
 PS Claim 1; Fig 6; 31pp; French.

XX mRNA was isolated from rabbit gastric mucosal cells, converted to
 CC cDNA and fragments cloned in pUC18. Recombinant plasmids were used
 CC to transform E.coli MW294 and cells screened for reactivity with
 CC probe corresp. to known parts of the RGL sequence. One positive
 CC clone contd. plasmid pJ0101 which contd. a 1.35 kb pscI sequence.
 CC This fragment was isolated, subjected to PCR and the amplification
 CC prod. ligated with the 2.2 kb BglII-SacI fragment of pRU276 (contg.
 CC a synthetic Tac promoter and transcription stop signal). The
 CC recombinant plasmid was designated prGLN2.1.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX

SQ Sequence 1378 BP; 397 A; 287 C; 276 G; 418 T; 0 other;

Query March 38.6%; Score 489.2; DB 14; Length 1378;
 Best Local Similarity 65.0%; Pred. No. 5.8e-145;
 Matches 722; Conservative 0; Mismatches 388; Indels 0; Gaps 0;

Qy 125 ACCGAGAGCATTTCATGATATATAGTGAATCATCCACATCAAGCTATCCCTGTGAGG 184
 Db 136 ACCCTGAAAGTGAATATGAATATATAGTCAGATGATTTCTTACTGGGATACCCAAGTGAA 195
 Qy 185 AATATCAAGTCGCAACTGAAGATGGGTATATCTTCTGTAAACAGAGTTCCTCCAGGCG 244
 Db 196 AATATGAGTTGTGACTGAAGATGGTATATCTTGAAGTAAACAGAAATCTCTATGGGA 255
 Qy 245 TAGTGCACCTAAGAGACAGGTTCCAGGCTGTGGTGTACTGACAGCATGGCTAGTTG 304
 Db 256 AGAAAAATTCAGGGAACAGAGACCCGCTGATATTTTTCAGCATGGTTGCTTG 315
 Qy 305 GAGGTGCTAGCACTGGATTTCCAACTGCCCAACATAGCTGGGCTTCATTCTGGCAG 364
 Db 316 CATCAGCCTCAAACTGGATCTCCAACTGCCCAACAGAGCTGGCTTTATCTTGGCAG 375
 Qy 365 ATGCTGGTTTGTGAGTGGGAAACAGAGGGGAAACCGCTGTCTCGAAACACA 424
 Db 376 ATGCTGGTTATGTTGTGTGCTGGGAAACAGAGAGAAATACCTGGTCCAGAGAAACT 435
 Qy 425 AGACACTCTCCATAGACCAAGATGAGTCTGGGCTTTTCACTATGATGAGTGGCTAGT 484
 Db 436 TATACATTTCCAGCACTCCGTTGAATCTGGCTTTTAGCTTTGATGAATGGCTAAAT 495
 Qy 485 TTGACCTTCTCGCAGTGATAAACTTTATTTTGCAGAAAAACGGGCCAGGAAAAAGATCTAT 544
 Db 496 ATGACCTTCCAGCCACATATGACTTCATTGTAAAGGAACTGGACAGGAAAAAGTCACT 555
 Qy 545 ATGTCGCTATTTCAGAGGACCAACATGGGCTTTATTTGCATTTTCCACCATGCCAGAGC 604
 Db 556 ATGTTGGTCAATTCAGGAGCACCAACATGGTTTATTTGCTCTTCTACAAATCCCAAGC 615
 Qy 605 TGGCTCAGAAAAATCAAAATGATTTTGTCTTAGCACCCTAGCCTGTTAAGCATGCA 664
 Db 616 TGGCTGAAGATCAAAACCTTTTATGATAGTCCAGTCCGCTGAGAGATACAA 675
 Qy 665 AAAGCCCCGGGACCAAAATTTTGTGCTGCCAGATATGATATCAAGGATTTGTTGGCA 724
 Db 676 AAAGCCTTGTAAACAAACTTAGGTTTATTTCTCCACCACTGTTTCAAGATATATTTGGTG 735
 Qy 725 AAAAGAAATTTCTGTATCAGCCAGAGATTTCTCAGACAACTGTTATTTTACCTTTGGGCC 784
 Db 736 ACAAATATTTCTACCCACACAAATTTCTTTGATCAATTTCTGCCACTCAAGTGTGTCCC 795
 Qy 785 AGGTGATTTCTTGATCAGATTTGTAGTAAATATCATGTTACTTCTGGGTGGATTCACACCA 844
 Db 796 GTGAGACACTGAAATGTAATTTGCAGCAATGCCCTTATTTATCATTTTGTGGCTTTGACAGG 855
 Qy 845 ACATATGAACATGAGCCGAGAGTGTATATGCTGCCACACTCTTGTCTGGAACATCTG 904
 Db 856 CAAACTGAACATGAGTCCGCTTGGATGTGTAGCTGTGCATATATCCGGCAGAACTTCAG 915
 Qy 905 TGCAAAATATTTACACATGAGCCAGGAGTGAATTTCTGTGTGAATCTCCGGGCAATTTGACT 964
 Db 916 TTCAAAACATGCTGCACTGGACCCAGGCTGTAAATCTGAAATTTTCAAGCTTTTAAAT 975

QY 1139 ATATTCCTGAATGGGCTCACGTGGATTTCATCTGGGGTTGGATGCTCTCACCGTATGT 1198
 Db 1143 AGATTCCTTTTACAAATCACTTGGACTTTATCTGGCAATGATGCCCTCAAGAGTTT 1202

QY 1199 ACAATGAATCATCCATCTGAT 1220
 Db 1203 ACAATGACATTTGTTCTATGAT 1224

RESULT 14

AAN60566

ID AAN60566 standard; DNA; 1367 BP.

XX

AC AAN60566;

XX 25-MAR-2003 (updated)

DT 22-AUG-1991 (first entry)

XX

XX Sequence encoding human pregastric lipase.

XX Cystic fibrosis therapy; enzyme; lipase deficiency; ss.

XX

XX Homo sapiens.

XX

FH Key Location/Qualifiers
 CDS 47..103

FT /*tag= a

FT mat_peptide 104..1243

FT /*tag= b

XX

XX WO8601532-A.

XX

XX 13-MAR-1986.

XX

XX 15-AUG-1985; 85WO-GB00364.

XX

XX 21-AUG-1984; 84GB-0021210.

XX

XX 15-AUG-1985; 85WO-GB00364.

XX

XX 01-JAN-1986; 86GB-000897.

XX

XX (CLLT) CELLTECH LTD.

XX

XX (LOWE/) LOWE P A.

XX

XX Lowe PA;

XX

XX WPI; 1986-081634/12.

XX

XX P-PSDB; AAP60658.

XX

XX New gastric lipase protein, esp. of human origin - for treating

XX lipase deficiency, and DNA sequences coding for it

XX

XX Disclosure; Fig 3; 39pp; English.

XX

XX The inventors claim a pregastric lipase protein and a gene encoding

XX it. Gastric lipase protein is useful for oral administration to

XX treat lipase deficiency, e.g. cystic fibrosis or pancreatitis.

XX (Updated on 25-MAR-2003 to correct PA field.)

XX

XX Sequence 1367 BP; 397 A; 284 C; 269 G; 417 T; 0 other;

XX

XX Query Match 37.0%; Score 470; DB 7; Length 1367;

XX Best Local Similarity 64.2%; Pred. No. 7.7e-139;

XX Matches 707; Conservative 0; Mismatches 395; Indels 0; Gaps 0;

XX

QY 119 CTGTGGACCCAGAGCATTCATGAATATTAGTGAATATCCACATCAAGGCTATCCCT 178

Db 123 CTGGAAGCCCTGAAGTGACTATGACATTAGTCAGATGATTACTTATTGGGGATACCCAA 182

QY 179 GTGAGGAATATGAAGTCGCAACTGAAGATGGGTATATCTTTCTGTTAAACAGATTTCCTC 238

Db 183 ATGAGGAATATGAAGTTGTGACTGAGAGATGGTATTATTTCTTGAAGTCAATAGAAATTCCTT 242

QY 239 GAGGCTTAGTGCACCACTAAGACAGACAGGTTCCAGGCTGTGGTGTGTACTGCAGCATGGCC 298
 Db 243 ATGGGAAGAAAATTACAGGAATACAGGCCAGAGACCTGTGTGTGTGTGTGTGTGTGTGT 302
 QY 299 TAGTTGAGAGTGTAGCACTGATTTCCAACTGCTCCACATAGATAGCTTGGGCTTCATTC 358
 Db 303 TGCTTGATCAGCCACAACTGATTTCCAACTGCTCCGGAACAACAGCCTTGGCTTCATTC 362
 QY 359 TGGCAGATGCTGCTTTTGAAGTGTGGATGGGGAACAGCAGGGGGAACCGCTGCTCTCGAA 418
 Db 363 TGGCAGATGCTGCTTTATGATGTGTGGCTGGCAACAGCAGAGGAACAACCTGGGCCAGAA 422
 QY 419 AACAAAGACACTCTCCATAGACCAAGATGAGTCTTGGGCTTTTCAGTTATGATGATGG 478
 Db 423 GAAACTTGTACTATTCCACAGATTCAGTTGAATTTCTGGGCTTTTCAGCTTTGATGAAATGG 482
 QY 479 CTAGGTTTGAACCTTCTCTGAGTGTATAAATTTTATTTTGCAGAAAACGGGCGAGAAAAGA 538
 Db 483 CTAAATATGACCTTCCAGCCACAATCGACTTCAATTTAAAGAAAATCTGGACAGACGAGC 542
 QY 539 TCTATTATGTCGGCTATTCAAGGCGACCACTATGGGCTTTTATTCATTTTCCACCATGC 598
 Db 543 TACACTATGTGGCCATTCCAGGCGACCACTATGGTTTTATTTGCTTTTCCACCAATC 602
 QY 599 CAGAGCTGCTCAGAAAATCAAAATGATTTTCTTGTAGCTTACCCATAGCACTGTTAAGC 658
 Db 603 CCAGCTGGCTAAAGAAATCAAAACCTTCTATGCTCTAGCTCTGTTGCTGCTGCTGAGT 662
 QY 659 ATGCAAAAAGCCCGGACCAAAATTTTGTGTGTCAGATATGATGATCAAGGATGT 718
 Db 663 ATACAAAAGCCTTATAAACAACACTTAGATTGCTTCTCAATCCCTCTTCAAGTTATAT 722
 QY 719 TTGCAAAAAGAAATTTCTGTATCAGACAGATTCTCAGACAACTGTTTATTTTACCTTT 778
 Db 723 TTGCTGACAAAATATTCTACCCACACAACTTCTTGTATCAATTTCTTGTACTGAAATGT 782
 QY 779 GTGGCCAGGTTGTTCTTGTATCAGATTGTTAGTAAATATCATGTTTCTTGGGTGGATTCA 838
 Db 783 GCTCCCGTGAGTGTGAATCTCTTTGAGCAATGCTTATTTATAATTTGTGGATTG 842
 QY 839 ACACCAACAATATGAACATCAGCGCAGCAAGTGTATGTCGCCACACCTTTGCTGGAA 898
 Db 843 ACAGTAAAGAACTTTAACACAGATCGCTTGGATGTATCTATCACATAATCCAGCAGAA 902
 QY 899 CATCTGTGCAAAATTTCTACACTGGAGCCAGCAGTGAATCTGTTGTTGTTGTTGTTGTT 958
 Db 903 CTCTGTTTCAAAACATGTTCCATTTGGACCCAGGCTGTTAAGTCTGGAAATTCAGACT 962
 QY 959 TTGACTGGGGAGTGAACCAAAAATCTGAAAAAATGCAATCAGCAACTCTCTGTAAAGT 1018
 Db 963 ATGACTGGGAAGCCAGTTTCAAGATAGGATGCACTATGATCAGTCCCAACTCCCTACT 1022
 QY 1019 ACAGAGTCAGATATGACGGTCCCTACAGCAATGTGGACAGAGGTGAGGATGAGCTTT 1078
 Db 1023 ACAATGTGACAGCCATGAATGTACCAATTCAGTGTGGAAACGGTGGCAAGGACCTTTGG 1082
 QY 1079 CAATCCAGAAAGCGTGAATAATGCTGCTCTGAGGTGACCAACCTCATCTACCAATAAGA 1138
 Db 1083 CTGACCCCAAGATGTTGGCTTTTGTCTCCAAAATCTCCCAATCTTATTACCAAGG 1142
 QY 1139 ATATTCTGAAATGGGCTCAGCTGATTTTCACTGGGCTTTGGATGCTCTCAACCGTATGT 1198
 Db 1143 AGATTCTTTTAAATCAATCACTTGGACTTATCTGGGCAATGATGCTCCCTCAAGAAAGTTT 1202
 QY 1199 ACAATGAATCATCCATCTGAT 1220
 Db 1203 ACAATGACATTTTCTATGAT 1224

RESULT 15

AAT58916

ID AAT58916 standard; DNA; 1367 BP.

XX

AC AAT58916;
 XX 19-AUG-1997 (first entry)
 XX Human gastric lipase coding sequence.
 XX Duodenal; gastric; lipase; transgenic; plant; recombinant; extract; ss;
 KW food; absorption; fat; pancreatic; cystic fibrosis; exocrine; dairy;
 KW hydrolysis; trans-esterification; substrate; enzyme; biofuel; human.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FH 47..1243
 FT /*tag= a
 FT /product= gastric lipase
 FT 47..103
 FT sig_peptide /*tag= b
 FT 104..1240
 FT mat_peptide /*tag= c
 XX WO9633277-A2.
 XX 24-OCT-1996.
 XX 19-APR-1996; 96WO-FR00606.
 XX 20-APR-1995; 95FR-0004754.
 XX (BIOC-) BIOCEM SA.
 XX (LJOU) INST RECH JOUVEINAL.
 XX Baudino S, Benicourt C, Cudrey C, Gruber V, Lenee P;
 FI Merot B;
 XX WPI; 1996-485783/48.
 DR P-PSDB; AAW09383.
 XX Recombinant human or canine pre-duodenal lipase prodn. in transgenic
 PT plants - useful for facilitating absorption of fat, as bio-catalysts
 PT and for prodn. of bio-fuel
 XX Claim 3; Fig 4; 130pp; French.
 XX This is the nucleotide sequence encoding the human pre-duodenal (i.e.
 CC gastric) lipase enzyme. The sequence can be used to generate transgenic
 CC plants producing recombinant lipase in an enzymatically active form.
 CC Alternatively bases encoding amino acids 20-23 or 20-73 (1-4 or 1-54 of
 CC the mature protein, respectively) can be deleted to form the derivatives
 CC designated Delta-4 or Delta-54 respectively. Plants, or their extracts,
 CC expressing the lipases or the truncated derivatives, can be used:
 CC (a) as pharmaceuticals or food to facilitate absorption of fat, either
 CC in healthy subjects or in patients with inadequate levels of
 CC gastric/pancreatic lipase e.g. those with cystic fibrosis or exocrine
 CC pancreatic insufficiency, in the elderly or in patients undergoing medical
 CC treatment which alters fat adsorption; (b) for performing industrial or
 CC agricultural reactions, e.g. in processing of fats or in the dairy
 CC industry, for hydrolysis or transesterification reactions, etc., where
 CC the plant material may provide both enzyme and substrate. The transgenic
 CC plants can also be used for biofuel production.
 XX Sequence 1367 BP; 397 A; 284 C; 269 G; 417 T; 0 other;
 SQ
 Query Match 37.0%; Score 470; DB 17; Length 1367;
 Best Local Similarity 64.2%; Pred. No. 7.7e-139;
 Matches 707; Conservative 0; Mismatches 395; Indels 0; Gaps 0;
 QY 119 CTGTGGACCCGAGCATTCTGATATATAGTGAATCATCCACATCAAGGCTATCCCT 178
 DB |||||
 DB 123 CTGGAAGCCCTGAAGTACTATGACATTAAGTACATGATTTATTTGGGATACCCAA 182
 QY 179 GTGAGGAATATGAATCGCACTGAAGATGGGTATATCTTCTGTTAAACAGAGATTCCTC 238
 |||||

Search completed: February 19, 2004, 05:06:03
 Job time : 356.095 secs

DB 183 ATGAAGAATATGAAGTTGTGATGAAGATGTTATATTTCTTGAAGTCAATAGATTCCTT 242
 QY 239 GAGGCTAGTGCACACCTAAGAGACAGGTTCCAGGCTGTGGTGTGTACTCAGCATGGCC 298
 DB 243 ATGGGAGAAAATTCAGGGAATACAGGCCAGAGACCTGTGTGTTTTCAGCATGGTT 302
 QY 299 TAGTTGAGGTGTAGCAACTGGATTTCCAACTGCCCAACAATAGCTGGCTTCAATTC 358
 DB 303 TGCTTGATCATGCCACAACTGGATTTCCAACTGCCCAACAATAGCTGGCTTCAATTC 362
 QY 359 TGGCAGATGCTGGTTTGAAGTGGGATGGGAAACAGCAGGGAAGCCCTGGTCTCGAA 418
 DB 363 TGGCAGATGCTGGTTTGAAGTGGGATGGGAAACAGCAGGGAAGCCCTGGGCGAGAA 422
 QY 419 AACACAAGACACTCTCCATAGACCAAGATGATGTTCTGGGCTTTTCAGTTATGATGATGG 478
 DB 423 GAAACTTGTACTATTACCCAGATTCAGTTGAATTTCTGGGCTTTTCAGCTTTGATGAAATGG 482
 QY 479 CTAGGTTTGACCTTCCCTGAGTGAATAACTTTATTTGCAGAAAACGGCGCAGGAAAGA 538
 DB 483 CTAATATGACCTTCCAGCCCAATTCGACTTCAITTTGAAGAAAACCTGGACAGAGCAGC 542
 QY 539 TCTATTATGTCGGCTATTTCACAGGCGCACCATGGGCTTTTATTGCAATTTTCCACCATGC 598
 DB 543 TACACTATGTTGGCCATTTCCAGGCGCACCATGGGCTTTTATTGCTTTTCCACCATGC 602
 QY 599 CAGAGCTGGCTCAGAAAATCAAAATGATTTTGGCTTTAGCACCCATAGCCACTGTTAAGC 658
 DB 603 CAGGCTGGCTAAAAGAAATCAAAACCTTCTATGCTCTAGCTCCTGTTGGCCACTGTGAAT 662
 QY 659 ATGCAAAAAGCCCGGAGCACCATAATTTTGTGTCGCGAGATATGATGATCAAGGATGT 718
 DB 663 ATACAAAAGCCCTTATAAACAACCTAGATTGTTGTTCTCAATCCCTCTTCAAGTTTATAT 722
 QY 719 TTGGCAAAAAGAAATTTCTGATCAGATTTGTAGTAATATATCATGTTACTTCTGGTGGATTC 838
 DB 723 TTGCTGACAAAATATTTTACCCACACAACTTTCTTTGATCAATTTCTTGTCTACTGAAGTGT 782
 QY 779 GTGGCCAGCTGATCTTCTGATCAGATTTGTAGTAATATATCATGTTACTTCTGGTGGATTC 838
 DB 793 GCTCCGCTGAGATGCTGATCTCTTTGACGCAATGCTTATTTATTTATTTGATGATTTG 842
 QY 839 ACACCAACATATGAACATGAGCGCAGCAAGTGTATGTCGCCCAACACTCTTCTCTGAA 898
 DB 843 ACAGTAAGAACTTTAAACAGGAGTGGTGTATCTATCATATAATCCAGCAGGAA 902
 QY 899 CATCTGTCCAAAATTTCTACACTGGAGCAGGAGTGAATTTCTGTTGAACTCCGGGCTAT 958
 DB 903 CTTCGTTTCAAAAATGTTTCCATTTGACCCAGGCTGTTAGTCTGGGAAATTCAGACTTT 962
 QY 959 TTGACTGGGGAGTGGAGACCAAAAATCTGHAATAATGCAATCAGCCAACTCTCTGTAAGT 1018
 DB 963 ATGACTGGGAGGCCAGTTCAGAAATAGATGTCATGATGATGATGATGATGATGATGAT 1022
 QY 1019 ACAGAGTCAGAGATATGAGTCCCTCAGCAATGTGGAAGAGGAGTTCAGGAGTTCGCTTT 1078
 DB 1023 ACAATGTGACAGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1082
 QY 1079 CAATATCCAGAGACGCTGAAAATGCTGCTCTGAGGTGACCAACCTCATCTTACCATAGA 1138
 DB 1083 CTGACCCCAAGATGTTGGGCTTTTGTCTTCAAACTCCCAATCTTATTTTACCAAGG 1142
 QY 1139 ATATTCCTGAATGGGCTCAGTGGATTTTCACTCTGGGTTTGGATGCTCTCTCAGCGTATGT 1198
 DB 1143 AGATTCCTTTTACAAATCAGTGGACTTTATCTGGGCAATGATGATGATGATGATGATGAT 1202
 QY 1199 ACAATGAATATCATCCATCTGAT 1220
 DB 1203 ACAATGACATTTGTTCTATGAT 1224

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 04:49:02 ; Search time 78.5225 Seconds
(without alignments)
7133.182 Million cell updates/sec

Title: US-10-042-431-46

Perfect score: 1269

Sequence: 1 atgttggaacctgtctcaag.....gacgggtgagccgtattg 1269

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgm2_6/ptodata/2/ina/5A_COMB.seq.*
- 2: /cgm2_6/ptodata/2/ina/5B_COMB.seq.*
- 3: /cgm2_6/ptodata/2/ina/6A_COMB.seq.*
- 4: /cgm2_6/ptodata/2/ina/6B_COMB.seq.*
- 5: /cgm2_6/ptodata/2/ina/PCTUS_COMB.seq.*
- 6: /cgm2_6/ptodata/2/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	470	37.0	1367	4	US-09-348-930A-8
2	456.8	36.0	1411	4	US-09-186-489-3
3	456.8	36.0	1411	4	US-10-043-665B-3
4	453.6	35.7	1134	4	US-09-186-489-4
5	453.6	35.7	1134	4	US-10-043-665B-4
6	450.4	35.5	1134	4	US-09-186-489-1
7	450.4	35.5	1134	4	US-10-043-665B-1
8	449.8	35.4	8324	4	US-09-186-489-5
9	449.8	35.4	8324	4	US-10-043-665B-5
10	444.2	35.0	1137	1	US-08-227-108-2
11	444.2	35.0	1137	2	US-09-073-674-2
12	444.2	35.0	1140	1	US-08-227-108-4
13	444.2	35.0	1140	2	US-09-073-674-4
14	444.2	35.0	1146	1	US-08-227-108-6
15	444.2	35.0	1146	2	US-09-073-674-6
16	444.2	35.0	1198	4	US-09-348-930A-5
17	444.2	35.0	1528	2	US-08-227-108-1
18	444.2	35.0	1528	2	US-09-073-674-1
19	444.2	35.0	1528	4	US-09-348-930A-1
20	444.2	35.0	1528	4	US-09-348-930A-7
21	414.2	32.6	1360	4	US-09-820-001-1
22	388.8	30.6	1048	4	US-09-348-930A-3
23	104.2	8.2	22067	4	US-09-820-001-3
24	48.2	3.8	178	1	US-08-751-782-2
25	48.2	3.8	178	2	US-08-925-171-2
26	46.4	3.7	43676	3	US-09-356-952-12
27	38.6	3.0	1664976	4	US-08-916-421B-1

28	36.6	2.9	3271	2	US-08-852-806-1	Sequence 1, Appli
29	36.6	2.9	3271	3	US-09-163-669-1	Sequence 1, Appli
30	36.6	2.9	3282	1	US-08-276-852-154	Sequence 154, App
31	36.6	2.9	3282	1	US-08-276-852-154	Sequence 169, App
32	36.6	2.9	3282	1	US-08-899-575-154	Sequence 154, App
33	36.6	2.9	3282	1	US-08-899-575-154	Sequence 169, App
34	36.6	2.9	3282	1	US-08-899-575-154	Sequence 154, App
35	36.6	2.9	3282	1	US-08-899-575-154	Sequence 169, App
36	36.6	2.9	3282	5	PCT-US95-08743-154	Sequence 154, App
37	36.6	2.9	3282	5	PCT-US95-08743-154	Sequence 169, App
38	36.6	2.9	13254	1	US-08-276-852-156	Sequence 156, App
39	36.6	2.9	13254	1	US-08-276-852-170	Sequence 170, App
40	36.6	2.9	13254	1	US-08-899-575-156	Sequence 156, App
41	36.6	2.9	13254	1	US-08-899-575-156	Sequence 170, App
42	36.6	2.9	13254	1	US-08-899-575-156	Sequence 156, App
43	36.6	2.9	13254	1	US-08-899-575-170	Sequence 170, App
44	36.6	2.9	13254	5	PCT-US95-08743-156	Sequence 156, App
45	36.6	2.9	13254	5	PCT-US95-08743-170	Sequence 170, App

ALIGNMENTS

RESULT 1
US-09-348-930A-8
; Sequence 8, Application US/09348930A
; Patent No. 6573431

GENERAL INFORMATION:
; APPLICANT: Lenee, P

; APPLICANT: Gruber, V

; APPLICANT: Baudino, S

; APPLICANT: Merot, B

; APPLICANT: Benicourt, C

; APPLICANT: Cudrey, C

; TITLE OF INVENTION: Recombinant Predudonal Lipases and Polypeptides Derivatives Prod

; TITLE OF INVENTION: Plants, Processes for Obtaining Them and Their Uses

; FILE REFERENCE: 18433/2012

; CURRENT APPLICATION NUMBER: US/09/348,930A

; CURRENT FILING DATE: 1999-07-02

; PRIOR APPLICATION NUMBER: 08/945,321

; PRIOR FILING DATE: 1998-02-12

; PRIOR APPLICATION NUMBER: PCT/FR96/00606

; PRIOR FILING DATE: 1996-03-19

; NUMBER OF SEQ ID NOS: 37

; SOFTWARE: Patent in version 3.0

; SEQ ID NO 8

; LENGTH: 1367

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (47)..(1240)

; US-09-348-930A-8

Query Match 37.0%; Score 470; DB 4; Length 1367;

Best Local Similarity 64.2%; Pred. No. 3.8e-142;

Mismatches 707; Conservative 0; Indels 0; Gaps 0;

QY	119	CTGTGACCCAGACGATTCATGATATAGTGAATATCATCCAAATCAAGCTATCCCT	178
Db	123	CTGGAAGCCCTGAAGTGACTATGAACATAGTCAGATGATTACTTATTTGGGATACCAA	182
QY	179	GTGAGGAATATGAAGTCGCACTGAAGATGGGTATATCTTTCTGTTAACAGGATTCCTC	238
Db	183	ATGGAATATGAAGTTGTGACTGAGATGGTTATATCTTGAATCAATAGATTCCTT	242
QY	239	GAGCCCTAGTGCACTTAAGAGACAGGTTCCAGGCTGTGGTGTACTGACGATGGCC	298
Db	243	ATGGGAAGAAAATTTCAGGGAATACAGGACCTGTGTGTTTTTGCAGCATGTT	302
QY	299	TAGTTCGAGGTGTACCACTGATTTCCAACTGCCCAACAAATAGCTGGGCTTCATTC	358
Db	303	TGCTTGATCAGCCCAAACTGATTTCCAACTGCCCAACAAAGCTTGGCTTCATTC	362

QY 939 TTCTGGTGAATCCGGGCAATTTGACTGGGGAGTGCAGCAAAATCTGGAATAATGCCAA 998
 Db 961 GTCTGGGAAATCCAAAGCTTTTGACTGGGGAGCCTCACTTGAAGAACCTTAATGCAATATAA 1020
 QY 999 TCAGCAACTCTGTAAAGTACAGAGTGCAGAGATAGCGGTCCCTACAGCAATGTGGAC 1058
 Db 1021 TCAGCCCAACCTCCCATCTCAATTTAAACAGCCATGAATGCCAATTCAGTATGGAG 1080
 QY 1059 AGAGGTTCAGGACTGGCTTCAATTCAGAGACAGTGAATGCTGCTCTCTGAGGTGAC 1118
 Db 1081 TCGTGGCCCAAGACCTGTGGCTGACCCCTCAGGATGTTGACCTTTTGGCTTTCAAAACTCTC 1140
 QY 1119 CAACCTCATCTACCAATAAGAAATTTCCGAATGGGCTCAGGTGGATTTTCACTGGGGTTT 1178
 Db 1141 TAATCTCATTCACCAAGAAATTTCCAAATTAACATCATCTGGACTTTATCTGGCAAT 1200
 QY 1179 GAATGCTCTCCAGTATGTAACAATGAATCATTCATCTGATGCAGAGGAGGA 1234
 Db 1201 GGATGACCTCAAGAAGTTTACAATGAATTTTCTTTTGTGGCAAAAGACAAA 1256

RESULT 3
 US-10-043-665B-3
 ; Sequence 3, Application US/10043665B
 ; Patent No. 6582948
 ; GENERAL INFORMATION:
 ; APPLICANT: Bolen, Paul L.
 ; APPLICANT: Cihak, Paul L.
 ; APPLICANT: Scharpf Jr., Lewis G.
 ; TITLE OF INVENTION: Recombinant Kid Pregastric Esterase and Methods for Its
 ; FILE OF INVENTION: Production and Use
 ; FILE REFERENCE: IFF-0009
 ; CURRENT APPLICATION NUMBER: US/10/043,665B
 ; PRIOR APPLICATION DATE: 2002-09-24
 ; PRIOR FILING DATE: 1998-11-05
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 1411
 ; TYPE: DNA
 ; ORGANISM: Kid (Goat)
 US-10-043-665B-3

Query Match 36.0%; Score 456.8; DB 4; Length 1411;
 Best Local Similarity 62.0%; Pred. No. 7.4e-138;
 Matches 741; Conservative 0; Mismatches 452; Indels 3; Gaps 1;
 QY 42 AATGGAATGGCTTCTGATTTCTGGTGGCGTATATGTTCCAGAGAAATGCAATTCAGT 101
 Db 61 ATTGAAATGGTGGCGTACTTGTAAACGGTGTGTTTCATCCACATGTCTGGAATGCCAT 120
 QY 102 ACATATGC---CAACTAAAGCTGTGGACCCAGAACATTCATGAATATTTAGTGAATCAT 158
 Db 121 TTGTTTCTTGGAAATTTGCTAAGAACCTTGAAGCCAGTATGAATGTCAGTATGAT 180
 QY 159 CCACATCAAGGTATCCCTGTGAGGAATATGAAGTCGCAACTGAAGATGGGTATATCCT 218
 Db 181 TTCTCTCTGGGCTACCCCAAGTGAATGATGATTAAGTTATTAAGTGAATGCTATATCCT 240
 QY 219 TTCTGTAAACAGATTCCTCGAGGCTAGTGAACCTAAGAACAGAGAGAGGTTCCAGGCTGT 278
 Db 241 TCAGGTCTATCGATTTCTTCATGGAAGAAATGATGCTATCATTTAGTTCAGAGACCTGT 300
 QY 279 GGTGTTTACTGCAGTATGGCTAGTTGGAGTGTCTAGCACTGGATTTTCAACCTGCCAA 338
 Db 301 TGTGTTTCTGCAGATGCTCTTCTCCCTCAGCTACAAACTGATTTTCAACCTTTCCAA 360
 QY 339 CAATAGCTGGGCTTCATTTCTGCAGATGCTGGTTTGGAGTGTGGATGGGGAACAGCAG 398
 Db 361 CAACAGCCTGGGCTTCTCTCTGTCAGATGCTGGTTATGACGTGTGGCTGGGGAACAGCAG 420

QY 399 GGGAAGCGCTGCTCGAAGAACAGACACTCCATAGACCAAGATGAGTCTGTGGC 458
 Db 421 AGGAACACTTGGGCCAGGAACATTTATCTATTTCCAGACTCCCTGGAATCTGGC 480
 QY 459 TTTCACTATGATGAGTGTGCTGCTTGTGACCTTCTGCAAGTATAAACTTTATTTGCA 518
 Db 481 TTTCACTATGATGAGTGTGCTGCTTGTGACCTTCTGCAAGTATAAACTTTATTTGCA 540
 QY 519 GAAACGGGCCAGGAAGATCTATTTATCTGCTGCTTATTCAGAGGACCAACCATGGCT 578
 Db 541 GAAACAGGACAGAGACTACATGTTGGCCATTTCCAGAGGACCAACCATGGCT 600
 QY 579 TATTCATTTTCCACCATGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 638
 Db 601 TGTGCTTTTCTACCAATCCACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 QY 639 ACCATAGCAGCTGTGATGATGCAAAAGCCCGGAGCAAAATTTTGTGCTGCGAGA 698
 Db 661 CCAGTGGCCACAGTGAAGCACCAGAGCCTGTTTAACTTTCACCTTATTTCTCA 720
 QY 699 TATGATGATCAAGGGATTTGTTGGCAAAAGAAATTTCTGATCAGACAGATTTCTCAG 758
 Db 721 CTTCCTTTCAAGATTAATTTGTTGTAACAAATGTTTACCCACACAAATTTTGAACA 780
 QY 759 ACACTTGTATTTTACCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 818
 Db 781 ATTTCTGTTGTTGAAAGTGTCTCTGCTGAGACACTGCTGCTGCTGCTGCTGCTGCT 840
 QY 819 GTTACTTCTGGTGGATTCACACCAACCAATATGACATGAGCCGAGCAAGTGTATATGC 878
 Db 841 GTTGGCAATTTACTGGAGTGCACAAATAAACTTCAACATGAGTGGCTTGAATGTATGT 900
 QY 879 TGCCCACTCTTGTGCTGGAACATCTCTGCAAAATATTTTACACTGAGCCAGCAGTGA 938
 Db 901 AGCACAATAATCCAGCAGGAGCTTCTGTTCAAAACATCTCTCCACTGGAGACAGCTATA 960
 QY 939 TTCTGGTGAATCCGGGCACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 998
 Db 961 GTCTGGAAATTTCCAACTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 QY 999 TCAGCAACTCTGTGAGTACAGAGTACAGATATGAGGCTCCCTACAGCAATGTGGAC 1058
 Db 1021 TCAGCCCACTCTCCCATCTACAAATTTACAGCCATGAATGTTCCCAATTCAGTATGGAG 1080
 QY 1059 AGAGGTTCAGACTGGCTTTCAATTCAGAGACGTGAATAATGCTGCTCTCTGAGGTGAC 1118
 Db 1081 TGCTGGCAAGACCTGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
 QY 1119 CAACCTCATCTACCAATAAGAAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1178
 Db 1141 TAATCTCATTCACCAAGGAATTTCAATTAACATCATCTGAGCTTTATCTGGCAAT 1200
 QY 1179 GGATGCTCTCAGCCTGATGTACAAATCAATCAATCAATCAATCAATCAATCAATCAAT 1234
 Db 1201 GGATGCACTCAAGAGTTTACAATGAATTTTCTTTTGTGGCAAAAGACAAA 1256

RESULT 4
 US-09-186-489-4
 ; Sequence 4, Application US/09186489
 ; Patent No. 6375947
 ; GENERAL INFORMATION:
 ; APPLICANT: Bolen, Paul L.
 ; APPLICANT: Cihak, Paul L.
 ; APPLICANT: Scharpf Jr., Lewis G.
 ; TITLE OF INVENTION: Purified Recombinant Kid Pregastric Esterase, and
 ; FILE OF INVENTION: Processes for its Production and Use
 ; FILE REFERENCE: 5499/3
 ; CURRENT APPLICATION NUMBER: US/09/186,489
 ; CURRENT FILING DATE: 1998-11-05
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 4

; LENGTH: 1134
 ; TYPE: DNA
 ; ORGANISM: Bovine
 ; US-09-186-489-4
 Query Match 35.7%; Score 453.6; DB 4; Length 1134;
 Best Local Similarity 62.7%; Pred. No. 7.1e-137;
 Matches 705; Conservative 0; Mismatches 419; Indels 0; Gaps 0;
 QY 1071 CTGGCTTTCAATCCAGAGACGTGAAATGCTGCTCTCTGAGGTGACCAACCTCATCTA 1130
 Db 969 CCTGTGGCTGACCCCTCAGGATGTTGACTTTCTGCTTCAAAACTCTCTAATCTCAATTA 1028
 QY 1131 CCATAGAGATATCTCTGAATGGGCTCAGTGGATTTCACTCTGGGTTTGGATGCTCTCA 1190
 Db 1029 CCACAAGAAATTCCTCAATTAATCAATCACTTTGGACTTTATCTGGCAATGGATGACCTCA 1088
 QY 1191 CCGTATGTACAAATCAATCAATCAATCTGATGACGAGGAGGAGA 1234
 Db 1089 AGAAGTTTACATGAATGTTCTTTGATGGCCGAGACAAA 1132
 RESULT 5
 US-10-043-665B-4
 ; Sequence 4, Application US/10043665B
 ; Patent No. 6582948
 ; GENERAL INFORMATION:
 ; APPLICANT: Boehr, Paul L.
 ; APPLICANT: Cihak, Paul, L.
 ; TITLE OF INVENTION: Recombinant Kid Prestagastic Esterase and Methods for Its
 ; FILE OF INVENTION: Production and Use
 ; FILE REFERENCE: IFP-0009
 ; CURRENT APPLICATION NUMBER: US/10/043,665B
 ; CURRENT FILING DATE: 2002-09-24
 ; PRIOR APPLICATION NUMBER: US 09/186,489
 ; PRIOR FILING DATE: 1998-11-05
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 4
 ; LENGTH: 1134
 ; TYPE: DNA
 ; ORGANISM: bovine
 ; US-10-043-665B-4
 Query Match 35.7%; Score 453.6; DB 4; Length 1134;
 Best Local Similarity 62.7%; Pred. No. 7.1e-137;
 Matches 705; Conservative 0; Mismatches 419; Indels 0; Gaps 0;
 QY 111 AACTAAAGCTGTGAGCCAGAGCAATTCATGAATTTAGTGAATCATCCACATCAAGG 170
 Db 9 AAAAATTCCTAAGAACCCCTGAAGCCAGTATGAATGTTAGTCAGATGATTTCTTACTCGGG 68
 QY 171 CTATCCCTGTGAGGAATATGAATGCGCACTGGAAGTGGGTATATCTTTCTGTAAACAG 230
 Db 69 CTACCAAGTGTGAGATGATGAATGTTAATCTGCGATGGTTATATCTTCAAGTCTATCG 128
 QY 231 GATTCCTCGAGCCCTAGTGCACCTTAAGAGACAGGTTCCAGGCTGTGGTGTACTGCA 290
 Db 129 GATTCCTCATGGAAGAAATATGCTATCATTTAGGTGACAGACCTGTGTGTCTTCTGCA 188
 QY 291 GCATGCGCTAGTGTGAGGTGCTAGCACTGATTTCCAACTGCTCCCACTATGCTCTGGG 350
 Db 189 GCATGGTCTTCTTGATCAGCCCAAACTGATTTCCAACTGCTCCCACTATGCTCTGGG 248
 QY 351 CTTCATTTCTGAGATGCTGTGTTTGTAGCTGTGATGGGGAACAGAGGGGAAACGCTG 410
 Db 249 CTTCCTCTGAGATGCTGTGTTATGACGTGTGGTGGGGAACAGAGAGAAACACCTG 308
 QY 411 GTCTGAAAAACACAGACACTCTCTCATAGACCAAGATGAGTCTGGGCTTTCAAGTATGA 470
 Db 309 GGGCCAGGAACTTTTATATCTATCCAGACTCCCGGAATTTCTGGGCTTTCAAGCTTTGA 368
 QY 471 TCAGATGGCTAGGTTTGAACCTTCTCTGCTAGTGTATTAATCTTTTTCAGAAAAACGGGCA 530
 Db 369 TGAATGGGGAATATGATGCTTCTCATCTTCAATTTGATTTTCAATTAAGAGACAGGACA 428
 QY 531 GGAAAGATCTATATGTCGGCTATTCACAGGCAACCACTGGGCTTTTATGATTTTC 590
 Db 429 GAAGAAGCTTACATATGTTGGCCATTTCCCAAGGCAACCACTGGGCTTTTATGCGCTTTTC 488
 QY 591 CACCATGCCAGAGCTGGCTCAGAAATCAGAAATGATTTTCTTTAGCACCATGACAC 650

; LENGTH: 1134
 ; TYPE: DNA
 ; ORGANISM: Bovine
 ; US-09-186-489-4
 Query Match 35.7%; Score 453.6; DB 4; Length 1134;
 Best Local Similarity 62.7%; Pred. No. 7.1e-137;
 Matches 705; Conservative 0; Mismatches 419; Indels 0; Gaps 0;
 QY 111 AACTAAAGCTGTGAGCCAGAGCAATTCATGAATTTAGTGAATCATCCACATCAAGG 170
 Db 9 AAAAATTCCTAAGAACCCCTGAAGCCAGTATGAATGTTAGTCAGATGATTTCTTACTCGGG 68
 QY 171 CTATCCCTGTGAGGAATATGAATGCGCACTGGAAGTGGGTATATCTTTCTGTAAACAG 230
 Db 69 CTACCAAGTGTGAGATGATGAATGTTAATCACTGCGATGGTTATATCTTCAAGTCTATCG 128
 QY 231 GATTCCTCGAGCCCTAGTGCACCTTAAGAGACAGGTTCCAGGCTGTGGTGTACTGCA 290
 Db 129 GATTCCTCATGGAAGAAATATGCTATCATTTAGGTGACAGACCTGTGTGTGTTCTGCA 188
 QY 291 GCATGCGCTAGTGTGAGGTGCTAGCACTGGAATTTCCAACTGCTCCCACTATGCTCTGGG 350
 Db 189 GCATGGTCTTCTTGATCAGCCCAAACTGATTTCCAACTGCTCCCACTATGCTCTGGG 248
 QY 351 CTTCATTTCTGAGATGCTGTGTTTGTAGCTGTGATGGGGAACAGAGGGGAAACGCTG 410
 Db 249 CTTCCTCTGAGATGCTGTGTTTGTAGCTGTGCTGGGGAACAGAGAGGAAACACCTG 308
 QY 411 GTCTGAAAAACACAGACACTCTCTCATAGACCAAGATGAGTCTGGGCTTTCAAGTATGA 470
 Db 309 GGGCCAGGAACTTTTATATCTATTCACAGACTCCCGGAATTTCTGGGCTTTCAAGCTTTGA 368
 QY 471 TGAGATGGCTAGGTTGACCTTCTGCTGAGTGAATTAATTTTTCAGAAAAACGGGCA 530
 Db 369 TGAATGGGGAATATGACCTTCTCATCTAATGATTTGATTTTCAATTAAGAGAGACAGGACA 428
 QY 531 GGAAAGATCTATATGTCGGCTATTCACAGGCAACCACTGGGCTTTTATGATTTTC 590
 Db 429 GAAGAAGCTTACATATGTTGGCCATTTCCCAAGGCAACCACTGGGCTTTTATGCGCTTTTC 488
 QY 591 CACCATGCCAGAGCTGGCTCAGAAATCAGAAATGATTTTCTTTAGCACCATGACAC 650
 Db 489 TACCAGTCCCACTATGGCTGAAAAATCAAGTCTTATGATGATAGCCCACTGTGCTAC 548
 QY 651 TGTTAAGATGCAAAAGCCCGGAGCAAAATTTTGTGTCGCGAGATGATGATCAA 710
 Db 549 AGTGAAGTACACCAAGAGCTGTTTAAACAACTTGCACTTATTCCTCACTTCTCTTCAA 608
 QY 711 GGGATTTGTCGCAAAAGAAATTTCTGTATCAGACAGAGATTTCTCAGACAACTTGTAT 770
 Db 609 GATTAATTTGTCGCAAAATGTTCTACCCACACACTTTTGGAACTTTCTTGGTGT 668
 QY 771 TTACCTTTGTCGAGGTGATTTCTGATGATTTGATGATTAATATCATGTTACTCTGGG 830
 Db 669 TGAATGTCCTCCGTCGAGACACTGGATGTCCTTTGTAAGAAATGCTTTGTCATATC 728
 QY 831 TGGATTTCAACCAACATATGAATGAGCCGAGCAAGTGTATATGCTGCCACACTCT 890
 Db 729 TGGAGTTGACAAATAAACTTCAACATGATGCTGCTTAGATGTTATATAGACATAATCC 788
 QY 891 TGCTGGAACATCTGTGCAAAATATTTCTACCTGAGCCAGGCACTGATGATTTCTGTGAAT 950
 Db 789 AGCAGGAATCTTGTTCAAAACACCCCTCACTGAGACAGGCTGTTAAGTCTGGAAAT 848
 QY 951 CCGGCAATTCAGTGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCACTCC 1010
 Db 849 CCAAGCTTTGACTGGGAGGCCCATATCAGAACTTAATGATTTATCATCAGGCCACAC 908
 QY 1011 TGTAAGGTACAGAGTACAGGATATGACGCTCCCTACAGCAATGTGGAGAGGAGGTGAGGA 1070
 Db 909 TCCCATCTACAAATTAAACAGCAATGATGTCCTCAATTTGCAATGAGTGTGACATGA 968

Db 489 TACCAAGTCCACATTTGGTGTAAATCAAGTCTTCTATGATAGCCAGTGGCAC 548
 Qy 651 TGTAAAGCATGCAAAAGCCGGGACCAAAATTTTGTGTGCGCAGATATGATGACAA 710
 Db 549 AGTGAAGTACACCAAGAGCGCTGTTTAAACAAATTTGCACTTATCTCCTCTTCAA 608
 Qy 711 GGGATTTGTTGGCAAAAGAAATTTCTATCAGACCAAGATTTCTCAGACAACTTGTAT 770
 Db 609 GATTAATTTGGTGCAAAATTTCTACCCACACACTTTTGGAACTTTCTTGTGT 668
 Qy 771 TTACCTTTTGGCCAGGTGATTTCTTGATCAGATTTGTAGTATATCATTTCTTCTGG 830
 Db 669 TGAATATGCTCCCTGAGACACTGATGCTCTTTGTAAAGTCTGTTGGCAATAC 728
 Qy 831 TGGATTCACACCAACATATCAACATCAGCGGAGCAGTGTATATGCTCCACACTCT 890
 Db 729 TGGATTTGACATAAAATCTCAACATGAGTCTTGTAGTGTATATAGCACATAATCC 788
 Qy 891 TGTGGAACATCTGTGCAAAATTTCTACACTGGAGCCAGGAGTGAATTTCTGGTGAAT 950
 Db 789 AGCAGAACTTCTGTTCAAAACACCCCTCCACTGGAGACAGGCTGTTAAGTCTGGAAAT 848
 Qy 951 CGGCAATTTGACTGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCCAACTCC 1010
 Db 849 CCAAGCTTTGACTGGGAGGCCATATCAGAACCTAATGCAATATCATCAGCCCAACC 908
 Qy 1011 TGTAAAGTACAGAGTACAGATATCAGCGTCTCCCTACAGCAATGTGGACAGGAGTCA 1070
 Db 909 TCCCATCTACAAATTTAAACGCCATGATGTCCCAATGTCAGTATGGAGTCTGACATGA 968
 Qy 1071 CTGGCTTTCAATCCAGAAAGCTGAAATGCTCTCTCAGGTGACCAACCTCATCTA 1130
 Db 969 CCGTGTGGCTGACCTCAGGAGTGTGACTTTCTGCTTTCAAAATCTCTAATCTCATTTA 1028
 Qy 1131 CCATAGAAATATCTCGAATGGCTCAGTGTGATTTCTATCTGGGTTTGGAGTCTCTCA 1190
 Db 1029 CCACAGGAATTCAAATATCAATCATCTGACCTTTATCTGGGCAATGGATGACCTCA 1088
 Qy 1191 CGATATGATCAATGAAATCATCCATCTGATGAGCAGGAGGAGA 1234
 Db 1089 AGAAGTTTACATGAAATGTTTCTTGTATGGCCGAGACAAAA 1132

RESULT 6

US-09-186-489-1
 ; Sequence 1, Application US/09186489
 ; Patent No. 6375947
 ; GENERAL INFORMATION:
 ; APPLICANT: Bolen, Paul L.
 ; APPLICANT: Chak, Paul L.
 ; APPLICANT: Scharpf Jr., Lewis G.
 ; TITLE OF INVENTION: Purified Recombinant Kid Pregastric Esterase, and
 ; TITLE OF INVENTION: Processes for its Production and Use
 ; FILE REFERENCE: 5499/3
 ; CURRENT APPLICATION NUMBER: US/09/186,489
 ; CURRENT FILING DATE: 1998-11-05
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 1134
 ; TYPE: DNA
 ; ORGANISM: Kid (Goat)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(1134)
 ; US-09-186-489-1
 Query Match 35.5%; Score 450.4; DB 4; Length 1134;
 Best Local Similarity 62.5%; Pred. No. 7,7e-136;
 Matches 703; Conservative 0; Indels 0; Gaps 0;
 Qy 111 AACTAAAGCTGGGACCCAGAGCAATTCATGAATATAGTGAATCATCCAAATCAAGG 170

Db 9 AAAAAATGCTAAGAACCCCTGAAGCCAGTATGAATGTGATCAGATGATTTCTCTCTGGG 68
 Qy 171 CTATCCCTGTGAGGAATATGAAGTGCACAACTCAAGATGAGGTATATCTTTCTGTGTAAACAG 230
 Db 69 CTACCAAGTGAAGTGCATATAAAGTTATTAAGTGCAGATGGCTATATCTTTCAGGTCTATCG 128
 Qy 231 GATTCCTCTGAGGCTAGTGCACAACTTAAGAGACAGGTTCCAGGCTGTGTGTCTTACTGCA 290
 Db 129 GATTCCTCATGGAAGAATGATGCTAATCATTTAGGTAGAGACCTGTGTGTGTCTGCA 188
 Qy 291 GCATGCCCTAGTGTGAGGCTGTAGCAACTGGAATTTCCAACTGCCCAACATATAGCCTGG 350
 Db 189 GCATGCTCTTCTGCTCAGCTACAACTGGAATTTCCAACTTCCCAACCAACAGCCTGG 248
 Qy 351 CTTCACTCTGGCAGATGCTGTTTGTGCTGTGGATGGGGAACAGCAGGGGAAACGCTGT 410
 Db 249 CTTCTCTCTGGCAGATGCTGTTTGTGCTGTGGGGAACAGCAGGGGAAACACTTG 308
 Qy 411 GTCTCGAAACCAAGACACTCTCCATAGACCAAGATGATTTCTGGGCTTTTCAGTTATGA 470
 Db 309 GGGCCAGGAACATTTATATCTATTTACCACTCCCTGAAATTCGGGCTTTTCAGCTTTGA 368
 Qy 471 TGAATGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 530
 Db 369 TGAATGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 428
 Qy 531 GGAAGATGCTATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 590
 Db 429 GAAGAGCTACTATGTTGGCCATTTCCAGGACCCACCATTTGTTGTGCTGCTTTC 488
 Qy 591 CACATGCGCAGAGTGGCTCAGAAATGATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 650
 Db 489 TACCAATCCACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 548
 Qy 651 TGTAAAGCATGCAAAAGCCCGGAGCAAAATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 710
 Db 549 AGTGAAGCACACCCAGAGCTGTTTAAACAACTTGCACTTATCTCTCTCTCTCTCTCTCT 608
 Qy 711 GGGATTTGTTGGCAAAAGAAATTTCTGATCAGACCAAGATTTCTCAGACAACTTGTAT 770
 Db 609 GATTAATTTGGTAAACAAATGTTCTACCCACACAAATTTTGTGAACAAATTTCTTGTGT 668
 Qy 771 TTACCTTTTGGCCAGGTGATTTCTTGATCAGATTTGTAGTATATCATTTCTTACTTCTGG 830
 Db 669 TGAAGTGTCTCTCTGAGACACTGATGCTCTTTTGAAGATGCTTGTGCTGCTGCTGCTGCT 728
 Qy 831 TGGATTCACACCAACATATGAACATGAGCCGAGCAAGTGTATATGCTGCCCACTCT 890
 Db 729 TGGAGCTGACATATAAACTTCAACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 788
 Qy 891 TGTGGAACATCTGTGCAAAATTTCTACACTGGAGCCAGGAGTGAATTTCTGTTGAAT 950
 Db 789 AGCAGAGCTCTGTTTCAAAACATCTCCACTGGAGACAGGCTATTTAAGTCTTGGGAAAT 848
 Qy 951 CGGCGATTTGACTGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCCAACTCC 1010
 Db 849 CCAAGCTTTGACTGGGAGCCCTCAGTGTGAACCTAATGCAATATATATCAGCCCAACC 908
 Qy 1011 TGTAAAGTACAGAGTACAGATATGAGGCTCCCTCAGCAATGTGGACAGGAGTCA 1070
 Db 909 TCCCATCTCAAAATTTAAACGCCATGATGCTCCCAATTTGAGTATGGAGTCTGGCAGAGA 968
 Qy 1071 CTGGCTTTCAATCCAGAGAGCTGGAATGCTGCTCTCTGAGGTGACCAACTCATCTA 1130
 Db 969 CTTGTTGGCTGACCTCAGGATGTTGACCTTTTGTCTTTTCAAACTCTCTAATCTCATTTCA 1028
 Qy 1131 CCATAAGAAATATTTCTGAAATGGGCTCACGTTGATTTTCATCTGGGTTTGGATGCTCTCTCA 1190
 Db 1029 CCACAGGAATTTCCAAATATCAATCATCTGACCTTTATCTGGCAATGATGACCTCA 1088
 Qy 1191 CCGTATGATCAATGAAATCATCCATCTGATGAGCAGGAGGAGA 1234

Db 1089 AGAAGTTTACATGAAATATTTCTTGTATGAGCAAGACAAA 1132

RESULT 7

US-10-043-665B-1

; Sequence 1, Application US/10043665B

; Patent No. 6582948

; GENERAL INFORMATION:

; APPLICANT: Bolen, Paul L.

; APPLICANT: Cihak, Paul, L.

; TITLE OF INVENTION: Recombinant Kid Pregastric Esterase and Methods for Its

; TITLE OF INVENTION: Production and Use

; FILE REFERENCE: IFF-0009

; CURRENT APPLICATION NUMBER: US/10/043,665B

; CURRENT FILING DATE: 2002-09-24

; PRIOR APPLICATION NUMBER: US 09/186,489

; PRIOR FILING DATE: 1998-11-05

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 1134

; TYPE: DNA

; ORGANISM: Kid (Goat)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(1134)

US-10-043-665B-1

Query Match 35.5%; Score 450.4; DB 4; Length 1134;

Best Local Similarity 62.5%; Pred. No. 7.7e-136; Indels 0; Gaps 0;

Matches 703; Conservative 0; Mismatches 421; Indels 0; Gaps 0;

Qy 111 AACTAAAGCTGTGGACCCAGAGCAATTCATGAATATTAGTGAATCATCAACATCAAGG 170

Db 9 AAAAATTCGTAAGAACCTCGAAGCCAGTATGAATGTGAGTCAGATGATTCCTTCCTGGG 68

Qy 171 CTATCCCTGTGAGGAATATGAAGTCGCAAGTGGGTATATATCTTCTTCTTAAAG 230

Db 69 CTACCAAGTGAGATGATGAAGTTAATCTGAGATGGGTATATCTTCTGAGTCTATCG 128

Qy 231 GATTCTCTGAGGCTTAGTGAACCTAAGAAAGACAGGTTCCAGGCTGTGTGTACTGCA 290

Db 129 GATTCTCTGAGGCTTAGTGAACCTAAGAAAGACAGGTTCCAGGCTGTGTGTACTGCA 188

Qy 291 GCATGGCTAGTTGGAGGTGTAGCACTGATTTCCAACTGCTGCCCAACATAGCTGGG 350

Db 189 GCATGGCTAGTTGGAGGTGTAGCACTGATTTCCAACTGCTGCCCAACATAGCTGGG 248

Qy 351 CTTCATCTGCGACATGCTGTTTGTGACGTGTGGATGGGAAACAGAGGGGAAACGCTG 410

Db 249 CTTCCTCTGCGACATGCTGTTTGTGACGTGTGGATGGGAAACAGAGGGGAAACGCTG 308

Qy 411 GTCTCGAAACACAGACATCTCTCCATAGACCAAGATGATTTCTGGGCTTTCAGTTAGA 470

Db 309 GGCCCGAGAACATTTATATCTATTCACCAAGTCTCCCTGGAATTTCTGGGCTTTCAGTTAGA 368

Qy 471 TGAGATGCTAGTTGACCTCTCTGAGTGATGAATTTTATTTGAGAAACGGGCCA 530

Db 369 TGAATGCTGATATGACCTTCTCATCTACATTTGATTTCTTTAAGAGAACAGACA 428

Qy 531 GGAAAAGATCTATTATGTGCGCTATTACAGGGCACCACCATGGCTTTTATGCAATTTTC 590

Db 429 GAAGAAGCTACACTATGTGTGGCCATTCCCAAGGCACCACCATGGTGTGTGTGCGCTTTTC 488

Qy 591 CACCATGCCAGAGTGTGCTGAGAAATCAAAATGATTTTGTCTTACGACCAATAGCCAC 650

Db 489 TACCAATCCCACTGTGCTGAGAAATTCGAAGTCTCCATGCAATAGCCCAATAGCCAC 548

Qy 651 TGTAAAGCATGCAAAAGCCCGGACCAAAATTTTGTGCTGCCAGATATGATGATCAA 710

Db 549 AGTGAAGCACACCCAGAGGCTGTGTTAACAACATTTGCACTTATTCCTCCTCTCTCAA 608

Qy 711 GGGATTTGTTGGCAAAAGAAATTTCTGTATCAGACCAAGATTTCTCAGACAACTGTAT 770

Db 609 GATTATATTTGGTAAACAAATGTTTACCCACACAAATTTTGTGAACAATTTCTGTGT 668

Qy 771 TTACCTTTTGGCCAGAGTATTTCTGTATCAGATTTGTAGTAATATCATGTTACTTCTGG 830

Db 669 TGAAGTGTCTCTCGTGAGACACTGGATGTCCTTTGTAAGAATGCGCTTTGTTGCCATTAC 728

Qy 831 TGGATTCAACACCAACAATATGAACATGAGCCGAGCAAGTGTATATGCTGCCACACTCT 890

Db 729 TGGAGCTGACATATTAACAACTTCAACATGAGTCGTTAGATGTGTATGATAGCAATATCC 768

Qy 891 TGCTGGAACTCTCTGCAAAATATTCTACACTGAGCCAGGCAAGTGAATTTCTGTTGAACT 950

Db 789 AGCAGGAGCTTCTCTTCAAAACATCTCCACTGGAGACAGGCTATTAAAGTCTGGGAAAT 848

Qy 951 CCGGCGATTTGACTGGGGAGTGAGACCAAAATCTGAAAAATGCAATCAGCAACTCC 1010

Db 849 CCAAGCTTTTGTACTGGGAGGCTCAGTTGAGACCTTAATGCAATATATATCAGCCACACC 908

Qy 1011 TGTAAAGTACAGAGTCAGAGATATGACGGTCCCTACAGCAATGTGGACAGAGTCAAGA 1070

Db 909 TCCCATCTACAATTTAAACAGCCATGAATGTCCCAATTCAGATGAGTGTGGCCAAAGA 968

Qy 1071 CTGGCTTTCAATCCAGAGAGCGTGAATGCTCTCTCTGAGGTGACCACTCATCTA 1130

Db 969 CTTGTGGCTGACCTCAGGATGTTGACCTTTTCTTTTCAAACTCTTAACTCTCATCTCA 1028

Qy 1131 CCATAAGAATATCTGAAATGGGCTCACGTGGATTTTCATCTGGGGTTTGGATGCTCTCA 1190

Db 1029 CCACAGGAAATTCCAAAATACATCATCTGGACTTTATCTGGGCAATGATGACCTCA 1088

Qy 1191 CCGTATGTACAAATGAATATCATCTGATGAGCAGGAGGAGA 1234

Db 1089 AGAAGTTTACAATGAATATTATTTCTTGTGGCAAAAGACAAA 1132

RESULT 8

US-09-186-489-5

; Sequence 5, Application US/09186489

; Patent No. 6375947

; GENERAL INFORMATION:

; APPLICANT: Bolen, Paul L.

; APPLICANT: Cihak, Paul L.

; APPLICANT: Scharpf Jr., Lewis G.

; TITLE OF INVENTION: Purified Recombinant Kid Pregastric Esterase, and

; TITLE OF INVENTION: Processes for its Production and Use

; FILE REFERENCE: 5499/3

; CURRENT APPLICATION NUMBER: US/09/186,489

; CURRENT FILING DATE: 1998-11-05

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 5

; LENGTH: 8324

; TYPE: DNA

; ORGANISM: Yeast YE-1 expression vector

US-09-186-489-5

Query Match 35.4%; Score 449.8; DB 4; Length 8324;

Best Local Similarity 62.7%; Pred. No. 3.9e-135; Indels 0; Gaps 0;

Matches 700; Conservative 0; Mismatches 417; Indels 0; Gaps 0;

Qy 118 GCTGTGGACCCAGAGCAATTCATGAATATTAGTGAATATCATCAACATCAAGGTATCCC 177

Db 1547 GCTAAGAACCCCTGAAGCCAGTATGAATGTGAGTCAGATGATTTCTTCTGGGGCTACCCA 1606

Qy 178 TGTGAGGAATATGAGTCGCACTGAGATGGGTATATCTTCTCTTAAACAGATTCCT 237

Db 1607 AGTGAGATGCAATAAGTTTAACTGAGATGGGTATATCTTCTTCAAGGTCTATCGGATTCCT 1666

Qy 238 CGAGGCTAGTGCAACCTAAGAACAGAGTTCAGAGGCTGTGGTGTACTGACAGATGGC 297

Db 1667 CATGGAAGAATGATGCTAATCATTTAGGTGAGAGACCTGTGTGTGTTCTTCAGCATGGT 1726

QY	298	CTAGTTGGAGGTGCTACGAACTGGAATTTCCAACTGCCCCAACAAATAGCCTGGGCTTCATT	357
Db	1727		
		CTTCTTCCCTCAGCTACAAACTGGATTTCCAACTTCCCAACCAACAGCCTGGGCTTCCTC	1786
QY	358	CTGGCAGATGCTGGTGTGCTGTGATGGGGAACAGCAGGGGAAACGCTGTGTCTCGA	417
Db	1787		
		CTGGCAGATGCTGGTGTATGACGTGTGCTGGGNAACAGCAGAGAAACAATTGGGCCAG	1846
QY	418	AAACACAAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTTCAGTTATGATGAGATG	477
Db	1847		
		GAACTTTATATATTACACAGACTCCCTCGAATTTCTGGGCTTTTCAGCTTTGATGAAATG	1906
QY	478	GCTAGGTTTGACCTTCTCGAGTGATAACTTTATTTTGCAGAAAACGGCCAGGAAAG	537
Db	1907		
		GCTGAAATGACCTTCCATCTCAATTTGATTTTCATCTTAAAGAAACAGGACAGAGAAG	1966
QY	538	ATCTATTATGTGGCTATTTCACAGGGCACCACTAGGCTTTATTTGCATTTTCCACCATTG	597
Db	1967		
		CTACACTATGTTGGCCATTCCCAAGGCACCACTATGTTTGTGCGCTTTTCTACCAAT	2026
QY	598	CCAGAGCTGGCTCAGAAAAACAAATGTATTTTGCCTTTAGACCCATAGCCACTGTTTAAG	657
Db	2027		
		CCCACTGGCTGAAAAAATGGAAGTCTTCCATGCATTAGCCCCAGTCGCCACAGTGAAG	2086
QY	658	CATGCAAAAAGCCCCGGGACCAAAATTTTGTGTGTCAGATATGATGATCAAGGGATTG	717
Db	2087		
		CACACCAGAGCGTTTTAACAACTTGCATTTACTCATCTCTCTTCTCAAGATTATTA	2146
QY	718	TTTGGCAAAAAGAAATTTCTGTATTCAGACAGAGTTTCTTCAGACAACCTGTGTATTTACCTT	777
Db	2147		
		TTTGGTAACAAAATGTTCTACCCACACAAATTTTGTGAAACAAATTTCTGGTGTGAAGTG	2206
QY	778	TGTGGCCAGGTGATTTCTTGATCAGATTTGTGTAGTAAATCATGTACTTCTGGGTGGATTC	837
Db	2207		
		TGCTCTCGTGAGACACGTGGATGTCCTTTGTAAGAAATGCCTTGTTGCGCATCTACTGGAGCT	2266
QY	838	AACACCAACATATGAACATGAGCGCAGCAAGTGATATGCTGCCACACACTCTTGCTGGGA	897
Db	2267		
		GACAAATAAAACTTCAACATGAGTCGCTTAGATGTGTATGTAGCACATAATCCAGCAGGA	2326
QY	898	ACATCTGTGCAAAATATTTACACTGAGCCAGCAGTGAAATTTGTGTGAATCTCCGGGCA	957
Db	2327		
		GCCTTCTGTTCAAAACATCTCCACTCGTGAGACAGGCTATTAAAGCTCGGGAAATTTCCAAGCT	2386
QY	958	TTTGACTGGGGAGTGAGACCAAAAATCTGGAATAATGCAATCAGGCCAACTCCTGTGAAG	1017
Db	2387		
		TTTGACTGGGGAGCGCTCAGTTGAGAACCTAATGTCAATTATTAATGAGCCACACTTCCCATC	2446
QY	1018	TACAGAGTCAGAGATATGACGGTCCCTACAGCAATGTGGACAGAGGTCAGGACTGGCTTT	1077
Db	2447		
		TACAATTTAACAGCCATGAATGTCCCAATTCAGTATGGAGTGTGCGCCAAAGACCTGTGTG	2506
QY	1078	TCAATCCAGNAGACGTGAAATGCTGCTCTCAGGTGACCACTCATCTACCATTAAG	1137
Db	2507		
		GCTGACCCCTCAGGATGTGACCTTTTGCTTTCAAACACTCTCTAATCTCATTCACCAAG	2566
QY	1138	AATATTCTGTAATGGGCTCAGTGGATTTTCATCTGGGGTTTGGATGCTCCTCACCCTGATG	1197
Db	2567		
		GAAATTTCCAAATTACAATCATCTGGACTTTATCTTGGGCAATGGATGCACCTCAAGAATTT	2626
QY	1198	TACAATGAATCATCCATCTGATGCAGCAGGAGAGA	1234
Db	2627	TACAATGAATTTATTTCTTTGATGGCAAAAGACAAAA	2663

RESIN.T 9

RESULT 9
US-10-043-665B-5

US-10-043-665B-5
; Sequence 5, Application US/10043665B

; sequence 3, Application
; Patent No. 6582948

GENERAL INFORMATION:

APPLICANT: Bolen, Paul L.

APPLICANT: Cihak, Paul, L.

Db 2327 GCCTCTGTTCAAAACATCTCCACTGGAGACAGGCTATTAAAGTCTGGAAATCCAGCT 2386
 Qy 958 TTTGACTGGGGAGTGAGACAAATAATCTGGAAAAATGCATCAGCCAACTCTCTGAAG 1017
 Db 2387 TTTGACTGGGGAGCTCAGTTGAGAACCTTAATGCATTAATCAGCCCACTCCCATC 2446
 Qy 1018 TACAGACTCAGAGATATGACGGTCCCTACAGCAATGTGGACAGAGGTGAGGACTGGCTT 1077
 Db 2447 TACAAATTACAGCCATGATGTCCTCAATTTGGAGTATGGTCTGGCCAGACCTGTTG 2506
 Qy 1078 TCAATCCAGAGAGCGTGAATAATGCTGCTCTCTGAGGTGACCACTCATCTACCAATAG 1137
 Db 2507 GCTGACCTCAGGATGTTGACCTTTTGCTTTCAAAACTCTCTAATCTCAATCACCACAAG 2566
 Qy 1138 AATATTCCTGAATGGGCTCAGGTGATTTTCATCTGGGTTTGGATGCTCTCACGGTATG 1197
 Db 2567 GAATTCCTCAATTAACAATCATCTGGACTTTATCTGGCAATGATGCACTCAAGATT 2626
 Qy 1198 TACAATGAATCATCCATCTGATGACAGCAGGAGAGA 1234
 Db 2627 TACAATGAATTAATTTCTTTGATGGCAAAAGACAAA 2663

RESULT 10

US-08-227-108-2
 ; Sequence 2, Application US/08227108
 ; Patent No. 5807726
 ; GENERAL INFORMATION:
 ; APPLICANT: Blanchard, Claire
 ; APPLICANT: Benicourt, Claude
 ; APPLICANT: Junien, Jean-Louis
 ; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/227,108
 ; FILING DATE: 03-APR-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fanucci, Allan A.
 ; REGISTRATION NUMBER: 30,256
 ; REFERENCE/DOCKET NUMBER: 7620-033
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212 790-9090
 ; TELEFAX: 212 869-8864/9741
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1137 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA to mRNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..1137
 ; US-08-227-108-2

Query Match 35.0%; Score 444.2; DB 1; Length 1137;
 Best Local Similarity 62.8%; Pred. No. 8e-134;
 Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

Qy 125 ACCCAGAGCATTCATGAATATTAGTGAATCATCAACATCAAGGCTATCCCTGTGAGG 184
 Db 26 ACCCTGAGTGACCATGATATATAGTCAATGATCACTACTGGGATACCCAGCTGAGG 85
 Qy 185 AATATGAAGTCGCAACTGAAGATGGGTATATCTCTTTCTGTAAACAGAGATTCTCGAGGCC 244
 Db 86 AATATGAAGTTGTGACCGAAGACGGTTATATCTCTGGGATCGACAGAAATTCCTTATGGGA 145
 Qy 245 TAGTGCACCTAAAGACACAGGTTCCAGGGCTGTGGTGTACTGCACATGGCTAGCTTG 304
 Db 146 GGAATAATTCAGAGAAATATAGGCGGAGACCTGTTGCATTTTTCACACAGGTTTGTCTG 205
 Qy 305 GAGGTGCTAGCAACTGGATTTCACACCTGCCCAACAAATAGCCTGGGCTTCACTTCTGSCAG 364
 Db 206 CATCAGCCACAACTGGATCTCCAACTGCCCAACACAGCCTGGCCCTTCATCTGCGCCG 265
 Qy 365 ATGCTGGTTTGTAGCTGTGGATGGGAAACAGCAGGGGAAAGCGCTGTCGAAAAACACA 424
 Db 266 AGCGCGGTACGACGTGTGGCTGGGAAACAGCAGGGGCAACACCTGGGCCAGGAGGATC 325
 Qy 425 AGACACTCTCCATAGACCAAGATGATTTCTGGGCTTTTCAGTTATGATGAGATGCTAGGT 484
 Db 326 TGTACTACTCGCCCGACTCCGTCGAATCTGGGCTTTTCAGCTTTGACGAGATGCTAAAT 385
 Qy 485 TTGACCTTCTGCTGAGTGAATAAATTTATTTTTCAGAAAAACGGGCCAGAAAAAGATCTATT 544
 Db 386 ATGACCTTTCCCGCCACCATTTGACTTTCATTTTGAAGAAAAACGGGACAGCAAGTACACT 445
 Qy 545 ATGCTGGCTATTTCACAGGGCACCACTGGGCTTTTATGCAATTTTCCACATGCGCAGC 604
 Db 446 AGTTGGCCANTTCCAGGGCACCACTGTTTCATCGCTTTTCCACCAATCCCAAGC 505
 Qy 605 TGGCTCAGAAAAATCAAAATGATTTTGTCTTTAGCACCCATAGCACCTGTTTAAGCATGCAA 664
 Db 506 TGGCGAAACGGATCAAAACCTTCTATGATAGCTCCGTTTCCACCGCTGAAGTACACCG 565
 Qy 665 AAAGCCCGGGACCAAAATTTTGTCTGCCAGATATGATGATCAAGGATTTGTTGGCA 724
 Db 566 AAACCTTGTAAACAAATCATGCTGCTCTTCTGCTCTTCAAGCTTATATTTGGAA 625
 Qy 725 AAAAAGAAATTTCTGATCAGACCAAGATTTCTCAGACAACTTTGTTATTTTACCTTTGGCC 784
 Db 626 ACATAATATTCTACCCACACCACTTCTTTGATCAATTTCTGCCACCGAGGTATGCTCCC 685
 Qy 785 AGGTGATTTCTGATCAGATTTGTAGTATATATGTTACTTCTGGTGGATTCAACACCA 844
 Db 686 GCGAGACGGTGGATCTCTCTGACGAAACGGCTGTTTATCATTTGTGGAATTTGACACTA 745
 Qy 845 ACAATATGAACATGAGCCGAGCAAGTGTATATGCTGCCCAACACTCTTGTCTGGAACATCTG 904
 Db 746 TGAATTTGACATGATGCTGCTTGGATGTATCTGTCAATATCCAGACAGAAACATCGG 805
 Qy 905 TCGAAAAATTTCTACACTGGAGCCAGGAGTGAATTTGTTGTAATCTGGGCACTTGGACT 964
 Db 806 TTCAGAACTGTCTCCACTGGTCCAGGCTGTTAAGTCTGGGAAGTTCACAGCTTTTGACT 865
 Qy 965 GGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCCAACTCTCTTAAGGTACAGAG 1024
 Db 866 GGGGAGGCCAGTTTCAGAAACATGATGCACTATCATCAGAGCATGCTCTCTACTACACC 925
 Qy 1025 TCAGAGATATGACGGTCCCTACAGCAATGTGGACAGGAGGTGAGGACTGGCTTTCAAATC 1084
 Db 926 TGACAGACATGATGCTGCAATCGCAGTGTGGAACGGTGGCAACGACATTTGCTGGCCGACC 985
 Qy 1085 CAGAGACGTGAAATGCTGCTCTGAGGTGACCAACCTCATCTACCATAGGATATTC 1144
 Db 986 CTCACGATTTGACCTTTTGTCTTCAAGCTCCCCAATCTCATTTTACCACAGGAGATTC 1045
 Qy 1145 CTGAATGGGCTCACGCTGGATTTTCACTGGGGTTTGGATGCTCTCACCGTATGTACAAATG 1204
 Db 1046 CTCCTTACATCATCTGGACTTTATCTGGGCCATGATGCCCCCTCAAGCGGTTTTACAAATG 1105

STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/227.108
FILING DATE: 03-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fanucci, Allan A.
REGISTRATION NUMBER: 30,256
REFERENCE/DOCKET NUMBER: 7620-033
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1140 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1140
US-08-227-108-4

Query Match 35.0%; Score 444.2; DB 1; Length 1140;
Best Local Similarity 62.8%; Pred. No. 8e-134;
Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

QY 125 ACCGAGACGATTCATGATATATAGTGAATATCATCAACATCAAGGCTATCCCTGTGAGG 184
DB 29 ACCCTGAAGTACCATGATATATAGTCAATGATCACTACTGGGATACCAGCTGAGG 88
QY 185 AATATGAAGTCGCAACTGAAGATGGTATATCCCTTTGTTAAACAGGATTCCTCGAGCC 244
DB 89 AATATGAAGTTGTGACCGAAGCGTTATATCCCTGGGATCGACAGAAATCCCTATGGGA 148
QY 245 TAGTCAACCTAAGAGACAGGTTCCAGGCTGTGGTACTGACATGCGCCTAGTTCG 304
DB 149 GGAATAATCAGAGATATAGCGGAGACCTGTTCATTTTGCACACGGTTGCTCG 208
QY 305 GAGGTGCTAGCAACTGGATTTCCAACTGCCCCAACAATAGCTGGGCTTCATTTCTGGCAG 364
DB 209 CATCAGCCACAACTGGATCTCCAACTGCCCAACAACAGCCTGSCCTTCATCTCTGGCCG 268
QY 365 ATGCTGGTTTTCAGCTGTGGATGGGAAACAGAGGGGAAACGCTGTCTCGAAACACA 424
DB 269 AGCGCGGTACAGCTGTGGTGGGGAACAGAGGGGCAACACCTGGGCGCAGGAGGAATC 328
QY 425 AGACACTCTCCATAGACCAAGATGATGTTCTGGGCTTTTCAGTTATGATGATGCTAGGT 484
DB 329 TGTAATCTCGCCGACCTCCGTAATTTCTGGCTTTTCAGTTTGAACGAGATGCTAAAT 388
QY 485 TTGACCTTCTGCACTGATAAATTTATTTTTCAGAAACCGGCGCAGGAAAGATCTATT 544
DB 389 ATGACCTTCCCGCCACCATTCATCTTGAAGAAACCGGACAGGACAGCAAGCTACACT 448
QY 545 ATGTCGGCTATTTCAGAGGCAACCATGGGCTTTATGCAATTTTCCACCATGCGAGGC 604
DB 449 ACCTTGGCCATTCACAGGCAACCATGTTTTCATGCTCTTTCCACCAATCCCAAGC 508
QY 605 TGGCTCAGAAATCAAAATGTTATTTTGTAGTCCCATAGCCACTGTTAAGCATGCAA 664
DB 509 TGGCGAAACGGATCAAAACCTTCTATGATGATGCTCCCGTTGCGCAACCGTGAAGTACACG 568
QY 665 AAGACCCCGGACCAAAATTTTGTGTCGCCAGATATGATCAAGGATGTTTGGCA 724

RESULT 13

US-09-073-674-4
Sequence 4, Application US/09073674
Patent No. 598189
GENERAL INFORMATION:
APPLICANT: Blanchard, Claire
APPLICANT: Benicourt, Claude
TITLE OF INVENTION: Recombinant Dog Gastric Lipase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Warner-Lambert Company
STREET: 2800 Plymouth Road
CITY: Ann Arbor
STATE: Michigan
COUNTRY: U.S.A.
ZIP: 48105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/073,674
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Crissey, Todd M.
REGISTRATION NUMBER: 37,807
REFERENCE/DOCKET NUMBER: 5072-D1-66-TMC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 734 622-7530
TELEFAX: 734 622-1553

86 AATATGAAGTTGTGACCGAGACGGTTATATCTTGGATCGACAGAAATTCCTTATGGGA 145
 245 TAGTGCACCTAAGAGACAGAGTTCCAGGCTGTGTGTACTGACGATGGCTAGTTG 304
 146 GGAATAATCAGAGATATAGGCGGAGACCTGTGCATTTTGGACACGGTTGCTCG 205
 305 GAGGTGTAGCACTGATTTCAACCTGCCCAACATAGCTGGGCTTCATTTGGCAG 364
 206 CATCAGCCACAACTGATCTCCAACTGCCCAACACAGCTGGCTTCATCTGGCCG 265
 365 ATGCTGTTTGTACGTGTGATGGGGAACAGCGGGGAAACGCTGTCTCGAAACACA 424
 266 ACGCGGGTACGACGTGTGGTGGGAAACAGCGGGGCAACACCTGGCCAGGAGATC 325
 425 AGACACTCTCCATAGACCAAGATGATTTCTGGGCTTCAGTTATGATGATGGCTAGT 484
 326 TGTAATCTCGCCGACTCGTGCATTTCTGGGCTTCAGTTTACGAGATGGCTAAT 385
 485 TTGACCTTCTGAGTGTAACTTTATTTTGGAGAAACGGGCGGAGGAAAGATCTATT 544
 386 ATGACCTTCCGCGCACCATTGACTTCTATCTTGAAGAAACGGGACAGGCAAGCTACACT 445
 545 ATGCGGCTATTCACAGGCGACCAACATGGGCTTTATGCAATTTCCACCTGCCAGGC 604
 446 ACGTTGGCCATTCACAGGCGACCAACATGGTTCATGCGCTTTTCCCAATTCCTCAAGC 505
 605 TGGCTCAGAAAAATCAAAATGATTTTCTTTAGCACCCATAGCACTGTTAAGCATGCAA 664
 506 TGGCGAAACGATCAAAACCTTCTATGATTTAGCTCCCGTTGCCACCGTGAAGTACACCG 565
 665 AAAGCCCGGGACCAATTTTGTGTCGCGCAATGATGATCAAGGATGTTTGGCA 724
 566 AAACCTTGTAAACAACTCATGCTCGTCCCTTCTGTTCTTCAAGCTTATATTTGGAA 625
 725 AAAAAGAAATTTCTATCAGACAGATTTCTCAGACAACTTGTATTTTACCTTTGGGC 784
 626 ACAAATTTCTACCCACACCACTTCTTGTATCAATTTCTGCGCACGAGGTATGCTCC 585
 785 AGGTGATTTCTGATCAGATTTGATGATTAATCATGTTACTTCTGGGTGATTCACACCA 844
 686 GCGAGACGGTGGATCTCTCTGACGCAACCGCTTGTATCATTTGTGATTTGACACTA 745
 845 ACAATATGAATGAGCGGAGCAAGTGTATGTCGCCACACTCTTGTCTGGAACATCTG 904
 746 TGAATTTGAACATGATGCTGTTGATGATGATCTGTCAATATCAGCAGGAAATCGG 805
 905 TGCAAAATATTTCTACTGAGCGCAGGAGTGAATTTCTGTTGAATCTCGGCGCATTTGACT 964
 806 TCCAGACGTGCTCCACTGTTCCAGGCTGTAACTCTGGGAAATTCCTCAAGCTTTTACT 865
 965 GGGGAGTGAGACCAAAATCTGAAATATGCAATGACCACTCTCTGTAAGTACAGAG 1024
 866 GGGGAGCCCAATTCAGAAATGATGATCATCATCAGAGCATGCTCTCCCTACTCAACC 925
 1025 TCAGATATGACCGTCTCTACAGCAATGTGGACAGGAGTCAAGGCTGCTTCAAAATC 1084
 926 TGAAGACATGATGTGCAATCGAGTGTGGACGCTGGCAAGCTTCTGGCCGACC 985
 1085 CAGAAGAGCGTGAATGCTGCTCTGAGGTGACCAACCTCATCTACATGAAATATTC 1144
 986 CTCAGATGTTGACCTTTGCTTTCCAAAGCTCCCAATCTCATTTACACAGGAAATTC 1045
 1145 CTGAATGGCTCAGTGTATTCATCTGGGGTTTGGATGCTCTCTCAAGTATGATCAATG 1204
 1046 CTCCTTACATCATTTGACCTTTATCTGGGCGCATGGATGCTTCAAGCGGTTTACATG 1105
 1205 AAATCATCATCTGATG 1221
 1106 AAATGTTTCCATGATG 1122

RESULT 15

US-09-073-674-6
 ; Sequence 6, Application US/09073674
 ; Patent No. 598189
 ; GENERAL INFORMATION:
 ; APPLICANT: Blanchard, Claire
 ; APPLICANT: Benicourt, Claude
 ; APPLICANT: Junien, Jean-Louis
 ; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Warner-Lambert Company
 ; STREET: 2800 Plymouth Road
 ; CITY: Ann Arbor
 ; STATE: Michigan
 ; COUNTRY: U.S.A.
 ; ZIP: 48105
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/073,674
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Crissey, Todd M.
 ; REGISTRATION NUMBER: 37,807
 ; REFERENCE/DOCKET NUMBER: 5072-D1-66-TMC
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 734 622-7530
 ; TELEFAX: 734 622-1553
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1146 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA to mRNA
 ; US-09-073-674-6

Query Match 35.0%; Score 444.2; DB 2; Length 1146;
 Best Local Similarity 62.8%; Pred. No. 8e-134;
 Matches 589; Conservative 0; Mismatches 408; Indels 0; Gaps 0;
 QY 125 ACCGAGAGCAATCATGAATATTAGTGAATATCCAAATCATCCAAATCAAGGCTATCCCTGTGAGG 184
 Db 26 ACCCTGAAGTGACCAATGAATATAAGTCAGATGATCACTACTGGGGATACCCAGCTGAGG 85
 QY 185 AATATGAAGTCGCAACTGAAGATGGGTATATCTCTTCTGTAAACAGGATTCCTCGAGGCC 244
 Db 86 AATATGAAGTTGTGACCGAGACGCTGTATATCTCTTGGATCGACAGAAATTCCTTATGGGA 145
 QY 245 TAGTGCAACCTAAGAAGACAGGTTCCAGGCGCTGTGGTGTACTGACAGCATGGGCTAGTTG 304
 Db 146 GGAATAATTCAGAGAAATATAGCGCGGAGACCTGTTCATTTTGGCAACACGCTTGTCTCG 205
 QY 305 GAGGTGTAGCAACTGGAATTTCCAACTGCCCAACATAGCTGGGCTTCATTTGCGAG 364
 Db 206 CATCAGCCACAACTGGATCTCCAACTGCCCAACACAGCTGGCTTCATCTGGCCG 265
 QY 365 ATGCTGTTTGTGACGTGTGGATGGGAAACAGCAGGGAACCGCTGTCTCGAAAAACACA 424
 Db 266 ACGCGGGTACGACGTGTGGCTGGGGAACACAGCGGGGCAACACCTGGCCAGGAGGATC 325
 QY 425 AGACACTCTCCATAGACCAAGATGATTTCTGGGCTTCAGTTATGATGATGGCTAGT 484
 Db 326 TGTACTACTCGCCGACTCCGTCGAAATTTCTGGGCTTTTCTGAGATGGCTAAAT 385
 QY 485 TTGACCTTCTGAGTGTAACTTTATTTTGGAGAAACGGGCGGAGGAAAGATCTATT 544
 Db 386 ATGACCTTCCGCGCACCATTGACTTCTTGAAGAAACGGGACAGGCAAGCTACACT 445

QY 545 ATGTGGCTATTACAGGCGACCAATGGGCTTTATTGATTTTCCACCATGCCAGAGC 604
 Db |||||
 QY 446 AGTTGGCCATTCAGGCGACCAATGGGCTTTATTGATTTTCCACCATGCCAGAGC 505
 Db |||||
 QY 605 TGGCTCAGAAATCAAAATGATTTTGTAGCAGCCATAGCCACTGTTAAGCATGCAA 664
 Db |||||
 QY 506 TGGCGAAACCGATCAAACTTCTATGATAGCTCCGCTTGCACCGTGAAGTACACCG 565
 Db |||||
 QY 665 AAAGCCCGGGACCAATTTTGTGCTGCCAGATATGATGATCAAGGGATTTGTTGCA 724
 Db |||||
 QY 566 AAACCTGTAAACAACTCATGCTGCTCCCTTCTCTTCAAGCTTATATTGGAA 625
 Db |||||
 QY 725 AAAAAGATTTCTGATCAGACCAATTTCTCAGACAACTTGTATTTTACCTTTGTGGCC 784
 Db |||||
 QY 626 ACAAATATTTTACCCACACCACTTCTTTGATCAATTTCTGCCACCGAGGTATGCTCCC 685
 Db |||||
 QY 785 AGGTGATTTCTGATCAGATTTCTGATCAATATATCATGTTACTTCTGGTGGATTCACACCA 844
 Db |||||
 QY 686 GCGAGCGGTGATTTCTCTCAGCAAGCCCTCTTTATCATTTGTGGATTTGACACTA 745
 Db |||||
 QY 845 ACAATATGAATGAGCGGAGCAAGTATATGCTGCCACACACTCTTCTGTAACATCTG 904
 Db |||||
 QY 746 TGAACCTTGAACATGAGTGGCTTGGATGTAATCTGACATAATCCACAGGAACTCGG 805
 Db |||||
 QY 905 TCGAAATATTTCTACTGAGCCAGGAGTGAATTTCTGTTGAATCTCGGGCATTTGACT 964
 Db |||||
 QY 806 TTCAGAACGTGCTCCACTGGTCCCAAGGCTTTAAGTCTGGGAAGTTCCAAAGCTTTGACT 865
 Db |||||
 QY 965 GGGGAGTGGAGACCAAAATCTGGAAATGCAATCAGCAACTCTCTGTAAGGTACAGAG 1024
 Db |||||
 QY 866 GGGGAAGCCCACTTCAGACATGATGCACTATCATCAGAGCATGCTCCCTACTACACC 925
 Db |||||
 QY 1025 TCAGAGATATGAGGTCCCTACAGCAATGTGGACAGGAGTCAAGACTGGCTTTCAAATC 1084
 Db |||||
 QY 926 TCAGACATGATGTCATGTCATCGCAATCGGAGTGTGGAACGTCGCAACGACTTGTGCGGACC 985
 Db |||||
 QY 1085 CAGAAAGCGTGAATGCTGCTCTGAGGTGACCAACTCATCTACCATAGATATTC 1144
 Db |||||
 QY 986 CTCACGATGTTGACCTTTTGTCTTCAAGCTCCCAATCTCATTTACCACAGGAGATTC 1045
 Db |||||
 QY 1145 CTGAATGGGCTCAGTGGATTTTCACTGGGGTTTGGATGCTCCTCACCGTATGTACAATG 1204
 Db |||||
 QY 1046 CTCCTTACAAATCACTTGGACTTTATCTGGGCCATGATGCCCTCAAGCGGTTTACAATG 1105
 Db |||||
 QY 1205 AAATCATCCATCTGATG 1221
 Db |||||
 QY 1106 AAATGTTTCCATGATG 1122
 Db |||||

Search completed: February 19, 2004, 10:33:59
 Job time : 84.5225 secs

GenCore version 5.1.6
Copyright (C) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 08:29:29 ; Search time 458.878 Seconds
(without alignments)
10186.874 Million cell updates/sec

Title: US-10-042-431-46

Perfect score: 1269

Sequence: 1 atgttggaacattgtcaag.....gacgtgtgagccgctattg 1269

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2449703 seqs, 1841816367 residues

Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/prodata/1/pubpna/PCT05_PUBCOMB.seq.*
- 7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/prodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/prodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq.*
- 16: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq.*
- 17: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq.*
- 18: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1269	100.0	1269	11	US-09-759-130B-416
2	1269	100.0	1269	14	US-10-042-431-46
3	1269	100.0	2044	11	US-09-759-130B-415
4	1269	100.0	2044	14	US-10-042-431-45
5	1267.4	99.9	2035	10	US-09-835-996A-20
6	1235.4	97.4	1384	10	US-09-835-996A-9
7	1122	88.4	1206	10	US-09-811-825-1
8	527.8	41.6	2481	10	US-09-880-107-3878
9	527.8	41.6	2858	15	US-10-198-846-13431
10	527.8	41.6	6952	13	US-10-133-013-144
11	525.4	41.4	2853	13	US-09-971-392-116
12	462.2	36.4	2506	12	US-10-108-260A-1262
13	456.8	36.0	1411	12	US-10-420-564-3
14	454.2	35.8	1194	13	US-10-312-088-7
15	453.6	35.7	1134	12	US-10-420-564-4

Sequence 3, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 5, Appli
Sequence 9, Appli
Sequence 31, Appli
Sequence 1, Appli
Sequence 27, Appli
Sequence 29, Appli
Sequence 6, Appli
Sequence 9462, Ap
Sequence 29, Appli
Sequence 4296, Ap
Sequence 6768, Ap
Sequence 4334, Ap
Sequence 4326, Ap
Sequence 4325, Ap
Sequence 1834, Ap
Sequence 1834, Ap
Sequence 253, App
Sequence 13032, A
Sequence 214253,
Sequence 214253,
Sequence 4162, Ap
Sequence 704, App
Sequence 3, Appli
Sequence 37615, A
Sequence 13508, A

ALIGNMENTS

RESULT 1

US-09-759-130B-416
; Sequence 416, Application US/09759130B
; Publication No. US2003002279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A
; APPLICANT: Fraser, Christopher C
; APPLICANT: Sharp, John D
; APPLICANT: Barnes, Thomas S
; APPLICANT: Kirt, Susan J
; APPLICANT: Mackay, Charles R
; APPLICANT: Myers, Paul S
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; TITLE OF INVENTION: USES.
; FILE REFERENCE: MP100-5350MIM
; CURRENT APPLICATION NUMBER: US/09/759,130B
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996

```

; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/602,871
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/420,707
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 416
; LENGTH: 1269
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-759-130B-416

Query Match      100.0%; Score 1269; DB 11; Length 1269;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGAACCTTGTCAAGACAGTGGATTTCTCACACAGATGGAAATGTGGCTTCG 60
DB 1 ATGTTGGAACCTTGTCAAGACAGTGGATTTCTCACACAGATGGAAATGTGGCTTCG 60
QY 61 ATTCTGGTGGCGTATATGTTCCAGAGAAATGTGAATTCAGTACATATGCCAACTAAAGCT 120
DB 61 ATTCTGGTGGCGTATATGTTCCAGAGAAATGTGAATTCAGTACATATGCCAACTAAAGCT 120
QY 121 GTGACCCAGAGCAATTCATGAATATTAGTGAATATCATCCAAATCAAGGCTATCCCTGT 180
DB 121 GTGACCCAGAGCAATTCATGAATATTAGTGAATATCATCCAAATCAAGGCTATCCCTGT 180
QY 181 GAGGAATATGAGTGCACCTGAAATGAGATGGGTATATCTTTCTGTTACAGATTCCTCGA 240
DB 181 GAGGAATATGAGTGCACCTGAAATGAGATGGGTATATCTTTCTGTTACAGATTCCTCGA 240
QY 241 GGCCTAGTGCACCTTAAAGACAGGTTCCAGGCTCTGGTGTATCTGACGATGGCCTA 300
DB 241 GGCCTAGTGCACCTTAAAGACAGGTTCCAGGCTCTGGTGTATCTGACGATGGCCTA 300
QY 301 GTTGGAGTGTACGACCTGATTTCCAACTGCTGCCACCAATAGCTGGGCTCATCTCG 360
DB 301 GTTGGAGTGTACGACCTGATTTCCAACTGCTGCCACCAATAGCTGGGCTCATCTCG 360
QY 361 GCAGATGCTGTTTGGAGTGTGATGGGGAACAGCAGGGGAAACGGCTGTCTCGAAAA 420
DB 361 GCAGATGCTGTTTGGAGTGTGATGGGGAACAGCAGGGGAAACGGCTGTCTCGAAAA 420
QY 421 CACAAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTTCAGTTATGATGATGGCT 480
DB 421 CACAAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTTCAGTTATGATGATGGCT 480
QY 481 AGGTTGACCTTCTCGAGTATATAAATTTATTTTGCAGAAAACGGGCCAGGAAAGATC 540
DB 481 AGGTTGACCTTCTCGAGTATATAAATTTATTTTGCAGAAAACGGGCCAGGAAAGATC 540
QY 541 TATTATGTCGGCTATTCACGGGACCAACCAATGGGCTTTATGCAATTTTCCACCATGCCA 600
DB 541 TATTATGTCGGCTATTCACGGGACCAACCAATGGGCTTTATGCAATTTTCCACCATGCCA 600
QY 601 GAGCTGGCTCAGAAATCAAAATGATTTTCTGCTTACGACCACTAGCCACTGTTAAGCAT 660
DB 601 GAGCTGGCTCAGAAATCAAAATGATTTTCTGCTTACGACCACTAGCCACTGTTAAGCAT 660
QY 661 GCAAAAAGCCCGGACCAAAATTTTGTGCTGCCAGATATGATGATCAAGGGATGTTT 720
DB 661 GCAAAAAGCCCGGACCAAAATTTTGTGCTGCCAGATATGATGATCAAGGGATGTTT 720
QY 721 GCAAAAAGATTTTCTGATCAGACGAGTTTCTCAGACAACTGTTATTTACCTTTGT 780
DB 721 GCAAAAAGATTTTCTGATCAGACGAGTTTCTCAGACAACTGTTATTTACCTTTGT 780
QY 781 GGGCAGGAGTATCTTGTATCAGATTTGTAGTATATCATGTTTCTGGGTGGATTCACAC 840
DB 781 GGGCAGGAGTATCTTGTATCAGATTTGTAGTATATCATGTTTCTGGGTGGATTCACAC 840

```

```

RESULT 2
US-10-042-431-46
; Sequence 46, Application US/10042431
; Publication No. US20020182675A1
; GENERAL INFORMATION:
; APPLICANT: MCCARTHY, Sean A.
; APPLICANT: BARNES, Thomas M.
; APPLICANT: FRASER, Christopher C.
; APPLICANT: SHARP, John D.
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
; TITLE OF INVENTION: PREVENTIVE, THERAPEUTIC, AND OTHER USES
; FILE REFERENCE: 10147-6U2
; CURRENT APPLICATION NUMBER: US/10/042,431
; CURRENT FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 46
; LENGTH: 1269
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-042-431-46

Query Match      100.0%; Score 1269; DB 14; Length 1269;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGAACCTTGTCAAGACAGTGGATTTCTCACACAGATGGAAATGTGGCTTCG 60
DB 1 ATGTTGGAACCTTGTCAAGACAGTGGATTTCTCACACAGATGGAAATGTGGCTTCG 60
QY 61 ATTCTGGTGGCGTATATGTTCCAGAGAAATGTGAATTCAGTACATATGCCAACTAAAGCT 120
DB 61 ATTCTGGTGGCGTATATGTTCCAGAGAAATGTGAATTCAGTACATATGCCAACTAAAGCT 120
QY 121 GTGACCCAGAGCAATTCATGAATATTAGTGAATATCATCCAAATCAAGGCTATCCCTGT 180

```

Db 121 GTGACCCAGAGCAATTCATGATATTTAGTGAATCATCCAAATCATCAAGGCTATCCCTGT 180
 Qy 181 GAGGAATATGAGTGGCACTGAAGATGGGTATATCCCTTCTGTAAACAGGATTCCTCGA 240
 Db 181 GAGGAATATGAGTGGCACTGAAGATGGGTATATCCCTTCTGTAAACAGGATTCCTCGA 240
 Qy 241 GGCCTAGTGCACCTTCAAGAGACAGAGTTCACAGGCTGTGGTGTACTGCACATGGGCTA 300
 Db 241 GGCCTAGTGCACCTTCAAGAGACAGAGTTCACAGGCTGTGGTGTACTGCACATGGGCTA 300
 Qy 301 GTTGGAGTGTAGCACTGAGTATCCAACTGCTCCCAATAGCTGGGCTTCATCTG 360
 Db 301 GTTGGAGTGTAGCACTGAGTATCCAACTGCTCCCAATAGCTGGGCTTCATCTG 360
 Qy 361 GAGATGCTGTTTTGAGTGTGGATGGGACACAGGGAAGAGCGCTGTCTCGAAA 420
 Db 361 GAGATGCTGTTTTGAGTGTGGATGGGACACAGGGAAGAGCGCTGTCTCGAAA 420
 Qy 421 CACAAGACACTCTCCATAGACCAAGATGAGTTCGGGCTTTCAGTATGATGAGATGGCT 480
 Db 421 CACAAGACACTCTCCATAGACCAAGATGAGTTCGGGCTTTCAGTATGATGAGATGGCT 480
 Qy 481 AGTTTACCTTCTGCTGAGTGAATTAACCTTTATTTGCAAGAAAACGGGCGAGAAAAGATC 540
 Db 481 AGTTTACCTTCTGCTGAGTGAATTAACCTTTATTTGCAAGAAAACGGGCGAGAAAAGATC 540
 Qy 541 TATTATGCTGGCTATTTCAGGGGACCAACCATGGCTTTATTTGCAATTTCCACCATGCCA 600
 Db 541 TATTATGCTGGCTATTTCAGGGGACCAACCATGGCTTTATTTGCAATTTCCACCATGCCA 600
 Qy 601 GAGTGGCTCAGAAAATCAAAATGATTTTGTCTTTCAGCACCATAGCCACTGTTAAGCAT 660
 Db 601 GAGTGGCTCAGAAAATCAAAATGATTTTGTCTTTCAGCACCATAGCCACTGTTAAGCAT 660
 Qy 661 GCAAAAGCCCGGACCAATTTTGTCTGCTGCAATATGATGATCAAGGATTTT 720
 Db 661 GCAAAAGCCCGGACCAATTTTGTCTGCTGCAATATGATGATCAAGGATTTT 720
 Qy 721 GCAAAAGCAATTTCTGTATCAGACCAAGATTTCTCAGACCACTTTTATTTACCTTTGT 780
 Db 721 GCAAAAGCAATTTCTGTATCAGACCAAGATTTCTCAGACCACTTTTATTTACCTTTGT 780
 Qy 781 GGCAGGTGATTTCTGTATCAGATTTGTAGTATATATCTTCTGCTGGTGGATTCAC 840
 Db 781 GGCAGGTGATTTCTGTATCAGATTTGTAGTATATATCTTCTGCTGGTGGATTCAC 840
 Qy 841 ACCAAATATGACATGACGCGACAGTGTATGCTGCCCACTCTTCTGCTGGAC 900
 Db 841 ACCAAATATGACATGACGCGACAGTGTATGCTGCCCACTCTTCTGCTGGAC 900
 Qy 901 TCTGTGCAAAATATTTCTACCTGAGCGCAGGAGTGAATTCGTGTGAATTCGGGCAATT 960
 Db 901 TCTGTGCAAAATATTTCTACCTGAGCGCAGGAGTGAATTCGTGTGAATTCGGGCAATT 960
 Qy 961 GACTGGGGAGTGACCAAAATCTGGAATATGCAATCAGCACTCTCTGTAAGTAC 1020
 Db 961 GACTGGGGAGTGACCAAAATCTGGAATATGCAATCAGCACTCTCTGTAAGTAC 1020
 Qy 1021 AGAGTCAGAGATATGACGCTCCCTACAGCAATGTGGAGGAGTGCAGGATGGCTTTCA 1080
 Db 1021 AGAGTCAGAGATATGACGCTCCCTACAGCAATGTGGAGGAGTGCAGGATGGCTTTCA 1080
 Qy 1081 AATCAGAGAGTGAATGCTCTCTGAGGTGACCAACCTCATCTACCAATGAAT 1140
 Db 1081 AATCAGAGAGTGAATGCTCTCTGAGGTGACCAACCTCATCTACCAATGAAT 1140
 Qy 1141 ATTCTGATGGGCTCAGTGAATTTCTCTGGGTTTGGATGCTCTCAGCGTATGTAC 1200
 Db 1141 ATTCTGATGGGCTCAGTGAATTTCTCTGGGTTTGGATGCTCTCAGCGTATGTAC 1200
 Qy 1201 AATGAAATCATCTATGATGACGAGGAGACCAACCTTTCCAGGAGCGGTGTAG 1260
 Db 1201 AATGAAATCATCTATGATGACGAGGAGACCAACCTTTCCAGGAGCGGTGTAG 1260

Qy 1261 GCGTATTG 1269
 Db 1261 GCGTATTG 1269

RESULT 3

US-09-759-130B-415
 ; Sequence 415, Application US/09759130B
 ; Publication No. US2003022279A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Millennium Pharmaceuticals, Inc.
 ; APPLICANT: McCarthy, Sean A
 ; APPLICANT: Fraser, Christopher C
 ; APPLICANT: Sharp, John D
 ; APPLICANT: Barnes, Thomas S
 ; APPLICANT: KIRST, Susan J
 ; APPLICANT: Mackay, Charles R
 ; APPLICANT: Myers, Paul S
 ; APPLICANT: Leiby, Kevin R
 ; APPLICANT: Wrighton, Nicolas
 ; APPLICANT: Goodearl, Andrew
 ; APPLICANT: Holtzman, Douglas A
 ; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
 ; PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
 ; USES.
 ; FILE REFERENCE: MPI00-5350NM
 ; CURRENT APPLICATION NUMBER: US/09/759,130B
 ; CURRENT FILING DATE: 2002-09-16
 ; PRIOR APPLICATION NUMBER: US 09/479,249
 ; PRIOR FILING DATE: 2000-01-07
 ; PRIOR APPLICATION NUMBER: US 09/559,497
 ; PRIOR FILING DATE: 2000-04-27
 ; PRIOR APPLICATION NUMBER: US 09/578,063
 ; PRIOR FILING DATE: 2000-05-24
 ; PRIOR APPLICATION NUMBER: US 09/333,159
 ; PRIOR FILING DATE: 1999-06-14
 ; PRIOR APPLICATION NUMBER: US 09/596,194
 ; PRIOR FILING DATE: 2000-07-14
 ; PRIOR APPLICATION NUMBER: US 09/342,364
 ; PRIOR FILING DATE: 1999-06-29
 ; PRIOR APPLICATION NUMBER: US 09/602,871
 ; PRIOR FILING DATE: 2000-06-23
 ; PRIOR APPLICATION NUMBER: US 09/420,707
 ; PRIOR FILING DATE: 1999-10-19
 ; NUMBER OF SEQ ID NOS: 460
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 415
 ; LENGTH: 2044
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-759-130B-415

Query Match 100.0%; Score 1269; DB 11; Length 2044;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 ATGTTGGAACCTTGTCAAGACAGTGGATTGCTCACAGAAATGGAATTCAGTACATATGCCAACTAAAGCT 60
 Db 126 ATGTTGGAACCTTGTCAAGACAGTGGATTGCTCACAGAAATGGAATTCAGTACATATGCCAACTAAAGCT 185
 Qy 61 ATTCTGCTGGCGTATATGTTCCAGAGAAATGGAATTCAGTACATATGCCAACTAAAGCT 120
 Db 186 ATTCTGCTGGCGTATATGTTCCAGAGAAATGGAATTCAGTACATATGCCAACTAAAGCT 245
 Qy 121 GTGGAGCCAGAGCAATTCATGAATATAGTAAATCATCCAAATCATCAAGGCTATCCCTGT 180
 Db 246 GTGGAGCCAGAGCAATTCATGAATATAGTAAATCATCCAAATCATCAAGGCTATCCCTGT 305

181 GAGGAATATGAAGTGGGAACTGAAGATGGGTATATCTCTTCTGTTTAAAGGATTCCTCGA 240
 306 GAGGAATATGAAGTGGGAACTGAAGATGGGTATATCTCTTCTGTTTAAAGGATTCCTCGA 365
 241 GGCCTAGTCAACTAAGAGAGACAGGTTCCAGGCTGTGGTGTACTGAGCATGGCCTA 300
 366 GGCCTAGTCAACTAAGAGAGACAGGTTCCAGGCTGTGGTGTACTGAGCATGGCCTA 425
 301 GTTGGAGTCTAGCACTGGATTTCCAACTGCCCAACATAGCTGGGCTTCATTCTG 360
 426 GTTGGAGTCTAGCACTGGATTTCCAACTGCCCAACATAGCTGGGCTTCATTCTG 485
 361 GCAGATGCTGTTTGAAGTGGGAACTGAAGTGGGAACTGAAGTGGGAACTGAAGTGGGAA 420
 486 GCAGATGCTGTTTGAAGTGGGAACTGAAGTGGGAACTGAAGTGGGAACTGAAGTGGGAA 545
 421 CACAAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTTCAGTTATGATGAGTGGT 480
 546 CACAAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTTCAGTTATGATGAGTGGT 605
 481 AGTTTTCACCTCTCTGAGTGAATATTTTTCAGAGAAACGGGCAAGGCAAGGAAAGATC 540
 606 AGTTTTCACCTCTCTGAGTGAATATTTTTCAGAGAAACGGGCAAGGCAAGGAAAGATC 665
 541 TATTATGCTGGCTATTCACAGGCGCACCATGGCTTTTATTCGATTTTCCACCATGCCA 600
 666 TATTATGCTGGCTATTCACAGGCGCACCATGGCTTTTATTCGATTTTCCACCATGCCA 725
 601 GAGTGGCTCAGAAAATCAAAATGTTTTCCTTTAGCACCCTAGCCTAGTATGAGTAT 560
 726 GAGTGGCTCAGAAAATCAAAATGTTTTCCTTTAGCACCCTAGCCTAGTATGAGTAT 785
 661 GCAAAAAGCCCGGACCAAAATTTTCTGTCGACATATGATGATCAAGGATTTGTTT 720
 786 GCAAAAAGCCCGGACCAAAATTTTCTGTCGACATATGATGATCAAGGATTTGTTT 845
 721 GGCAAAAGAAATTTCTGTATCAGACAGATTTCTCAGACAACTGTTGTTATTTACCTTTGT 780
 846 GGCAAAAGAAATTTCTGTATCAGACAGATTTCTCAGACAACTGTTGTTATTTACCTTTGT 905
 781 GGCAGAGTGTATTTGATCAGATTTGATTAATATCATGTTTCTCTGGTGGATTTCAAC 840
 906 GGCAGAGTGTATTTGATCAGATTTGATTAATATCATGTTTCTCTGGTGGATTTCAAC 965
 841 ACCAACAATATGAACATGAGCGGAGCAAGTGTATATCTGCCCACTCTTGTGGAACA 900
 966 ACCAACAATATGAACATGAGCGGAGCAAGTGTATATCTGCCCACTCTTGTGGAACA 1025
 901 TCTGTGCAAAATATTTCTACATGAGCGGAGCAAGTGTATATCTGCCCACTCTTGTGGAACA 960
 1026 TCTGTGCAAAATATTTCTACATGAGCGGAGCAAGTGTATATCTGCCCACTCTTGTGGAACA 1085
 961 GACTGGGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCCAACTCTCTGTAAGGTAC 1020
 1086 GACTGGGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCCAACTCTCTGTAAGGTAC 1145
 1021 AGAGTCAGAGATATGACGGTCCCTACAGCAATGTGGAAGAGTCAAGTCAAGTCAAGTCA 1080
 1146 AGAGTCAGAGATATGACGGTCCCTACAGCAATGTGGAAGAGTCAAGTCAAGTCAAGTCA 1205
 1081 AATCCAGAGAGAGTGAATAATGCTGCTCTGAGGTGACCAACCTCTCTACCAATAAGAT 1140
 1206 AATCCAGAGAGAGTGAATAATGCTGCTCTGAGGTGACCAACCTCTCTACCAATAAGAT 1265
 1141 ATTCCTGAATGGGTCAAGTGAATTTCACTGGGGTTTGGATGCTCTCAACCGTATGTAC 1200
 1266 ATTCCTGAATGGGTCAAGTGAATTTCACTGGGGTTTGGATGCTCTCAACCGTATGTAC 1325
 1201 AATGAATATCCATCTCAGCAGCAGGAGGAGACCACTTTTCCAGGGAGCGGTGTAG 1260
 1326 AATGAATATCCATCTCAGCAGCAGGAGGAGACCACTTTTCCAGGGAGCGGTGTAG 1385
 1261 GCCGTATTG 1269

1386 GCCGTATTG 1394
 RESULT 4
 US-10-042-431-45
 ; Sequence 45, Application US/10042431
 ; Publication No. US20020182675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MCCARTHY, Sean A
 ; APPLICANT: BARNES, Thomas M
 ; APPLICANT: FRASER, Christopher C
 ; APPLICANT: SHARP, John D
 ; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
 ; TITLE OF INVENTION: PREVENTIVE, THERAPEUTIC, AND OTHER USES
 ; FILE REFERENCE: 10147-6U2
 ; CURRENT APPLICATION NUMBER: US/10/042,431
 ; CURRENT FILING DATE: 2001-10-25
 ; PRIOR APPLICATION NUMBER: US 09/333,159
 ; PRIOR FILING DATE: 1999-06-14
 ; PRIOR APPLICATION NUMBER: US 09/578,063
 ; PRIOR FILING DATE: 2000-05-24
 ; NUMBER OF SEQ ID NOS: 79
 ; SOFTWARE: Patent in Ver. 2.1
 ; SEQ ID NO 45
 ; LENGTH: 2044
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-042-431-45
 Query Match 100.0%; Score 1269; DB 14; Length 2044;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGTGGGAACTGTGCAAGACAGTGGATTTCTCACAAGATGGAAATGGGCTTCTG 60
 DB 126 ATGTGGGAACTGTGCAAGACAGTGGATTTCTCACAAGATGGAAATGGGCTTCTG 185
 QY 61 ATTTCTGGGGCTATATGTTCCAGAGAAATGTGAATTCAGTACATATGCAATGCAAGCT 120
 DB 186 ATTTCTGGGGCTATATGTTCCAGAGAAATGTGAATTCAGTACATATGCAATGCAAGCT 245
 QY 121 GTGGACCCAGAGCATTCATGAATATTAGTGAATATCCAAATCAAGGCTATCCCTGT 180
 DB 246 GTGGACCCAGAGCATTCATGAATATTAGTGAATATCCAAATCAAGGCTATCCCTGT 305
 QY 181 GAGGAATATGAAGTGGCAACTGGAAGTGGGTATATCTCTGTTAAACAGGATTCCTCGA 240
 DB 306 GAGGAATATGAAGTGGCAACTGGAAGTGGGTATATCTCTGTTAAACAGGATTCCTCGA 365
 QY 241 GGCCTAGTGGCAACTTGAAGACAGAGTTCAGGCTGTGGTGTACTGAGCATGGCCTA 300
 DB 366 GGCCTAGTGGCAACTTGAAGACAGAGTTCAGGCTGTGGTGTACTGAGCATGGCCTA 425
 QY 301 GTTGGAGTGTCTAGCACTGGATTTCCAACTGCCCAACATAGCTGGGCTTCATTCTG 360
 DB 426 GTTGGAGTGTCTAGCACTGGATTTCCAACTGCCCAACATAGCTGGGCTTCATTCTG 485
 QY 361 GCAGATGCTGTTTGAAGTGGGAACTGAAGTGGGAACTGAAGTGGGAACTGAAGTGGGAA 420
 DB 486 GCAGATGCTGTTTGAAGTGGGAACTGAAGTGGGAACTGAAGTGGGAACTGAAGTGGGAA 545
 QY 421 CACAAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTTCAGTTATGATGAGTGGT 480
 DB 546 CACAAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTTCAGTTATGATGAGTGGT 605
 QY 481 AGTTTTCACCTCTCTGAGTGAATATTTTTCAGAGAAACGGGCAAGGCAAGGAAAGATC 540
 DB 606 AGTTTTCACCTCTCTGAGTGAATATTTTTCAGAGAAACGGGCAAGGCAAGGAAAGATC 665
 QY 541 TATTATGCTGGCTATTCACAGGCGCACCATGGCTTTTATTCGATTTTCCACCATGCCA 600
 DB 666 TATTATGCTGGCTATTCACAGGCGCACCATGGCTTTTATTCGATTTTCCACCATGCCA 725

QY 601 GAGCTGGCTCAGAAATCAAAATGATTTTGGCTTTAGCACCATAGCCACTGTTAAGCAT 660
 Db 726 GAGCTGGCTCAGAAATCAAAATGATTTTGGCTTTAGCACCATAGCCACTGTTAAGCAT 785
 QY 661 GCAAAAAGCCCGGACCAAAATTTTGGTGGTCCAGATATGATGATCAAGGGATTTGTT 720
 Db 786 GCAAAAAGCCCGGACCAAAATTTTGGTGGTCCAGATATGATGATCAAGGGATTTGTT 845
 QY 721 GCAAAAAGATTTCTGATCAGACCAAGATTTCTCAGACAACTTGTATTACCTTTGT 780
 Db 846 GCAAAAAGATTTCTGATCAGACCAAGATTTCTCAGACAACTTGTATTACCTTTGT 905
 QY 781 GCGCAGGTGATTTCTGATCAGATTTGTTAGTAAATATCATGTTACTTCTGGTGGATTCAC 840
 Db 906 GCGCAGGTGATTTCTGATCAGATTTGTTAGTAAATATCATGTTACTTCTGGTGGATTCAC 965
 QY 841 ACCAACAATAGAAATGAGCGGACCAAGATTTGTTAGTAAATATCATGTTACTTCTGGTGGATTCAC 900
 Db 966 ACCAACAATAGAAATGAGCGGACCAAGATTTGTTAGTAAATATCATGTTACTTCTGGTGGATTCAC 1025
 QY 901 TCTGTGCAAAATATTTCTACCTGAGCGGACCAAGATTTGTTAGTAAATATCATGTTACTTCTGGTGGATTCAC 960
 Db 1026 TCTGTGCAAAATATTTCTACCTGAGCGGACCAAGATTTGTTAGTAAATATCATGTTACTTCTGGTGGATTCAC 1085
 QY 961 GACTGGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCCAACTCTCTGTAAAGTAC 1020
 Db 1086 GACTGGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCCAACTCTCTGTAAAGTAC 1145
 QY 1021 AGAGTCAGAGATATGACGGTCCCTACAGATATGAGACAGAGTCAAGACTGCTTTCA 1080
 Db 1146 AGAGTCAGAGATATGACGGTCCCTACAGATATGAGACAGAGTCAAGACTGCTTTCA 1205
 QY 1081 AATCAGAAAGAGTGAATAATGCTCTCTGAGGTGACCAACCTCATCTACCAATAAGAT 1140
 Db 1206 AATCAGAAAGAGTGAATAATGCTCTCTGAGGTGACCAACCTCATCTACCAATAAGAT 1265
 QY 1141 ATTCTGAATGGGCTCAGTGGAATTTATCTGGGGTTTGGATGCTCTCTACCGTATGTAC 1200
 Db 1266 ATTCTGAATGGGCTCAGTGGAATTTATCTGGGGTTTGGATGCTCTCTACCGTATGTAC 1325
 QY 1201 AATGAAATCATCATCTGATGACGAGGAGGACCAACCTTCCAGGAGCGGTGTAG 1260
 Db 1326 AATGAAATCATCATCTGATGACGAGGAGGACCAACCTTCCAGGAGCGGTGTAG 1385
 QY 1261 GCGGTATTG 1269
 Db 1386 GCGGTATTG 1394

RESULT 5
 US-09-835-996A-20
 ; Sequence 20, Application US/09835996A
 ; Patent No. US20020142953A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ballinger, Dennis
 ; APPLICANT: Loeb, Debra
 ; APPLICANT: Montgomery, Julie
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Zhou, Ping
 ; APPLICANT: Goodrich, Ryle
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Zhao, Qing
 ; APPLICANT: Wehrman, Tom
 ; APPLICANT: Drmanac, Radoje
 ; APPLICANT: Ren, Feiyan
 ; APPLICANT: Qian, Xiaohong
 ; APPLICANT: Wang, Dunrui
 ; TITLE OF INVENTION: MATERIALS AND METHODS RELATING TO LIPID METABOLISM
 ; FILE REFERENCE: 28110/35915A
 ; CURRENT APPLICATION NUMBER: US/09/835,996A
 ; CURRENT FILING DATE: 2001-04-16

; PRIOR APPLICATION NUMBER: US 60/197,137
 ; PRIOR FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: US 09/714,936
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: US 09/667,298
 ; PRIOR FILING DATE: 2000-09-22
 ; PRIOR APPLICATION NUMBER: US 09/631,451
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: US 09/598,042
 ; PRIOR FILING DATE: 2000-06-20
 ; NUMBER OF SEQ ID NOS: 45
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 20
 ; LENGTH: 2035
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (189)..(1415)
 ; US-09-835-996A-20

Query Match 99.9%; Score 1267.4; DB 10; Length 2035;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1268; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATCTTGGAAACCTTGTCAAGACAGTGGATTGTCTCACACAGAAATGGAATGTGGCTTCTG 60
 Db 147 ATCTTGGAAACCTTGTCAAGACAGTGGATTGTCTCACACAGAAATGGAATGTGGCTTCTG 206
 QY 61 ATTCTGGTGGCTATATGTTCCAGAGAAATGGAATTCAGTACATATGCCAACTAAAGCT 120
 Db 207 ATTCTGGTGGCTATATGTTCCAGAGAAATGGAATTCAGTACATATGCCAACTAAAGCT 266
 QY 121 GTGGACCCAGAAAGCATTCATGAATATTAGTAAATCATCCAAATCATCAAGCTATCCCTGT 180
 Db 267 GTGGACCCAGAAAGCATTCATGAATATTAGTAAATCATCCAAATCATCAAGCTATCCCTGT 326
 QY 181 GAGGAATATGAATGCGAACTGAAAGTGGTATATCTCTTCTGTAAACAGGAATCTCTCGA 240
 Db 327 GAGGAATATGAATGCGAACTGAAAGTGGTATATCTCTTCTGTAAACAGGAATCTCTCGA 386
 QY 241 GGCTTAGTGCACCTAAGAGACAGTTCACAGGCTTGTGGTGTACTGTGAGCATGCGCTTA 300
 Db 387 GGCTTAGTGCACCTAAGAGACAGTTCACAGGCTTGTGGTGTACTGTGAGCATGCGCTTA 446
 QY 301 GTTGGAGGTGCTAGCAACTGGATTTCACACCTGCCCAACAATAGCTGGCTTCAATTCG 360
 Db 447 GTTGGAGGTGCTAGCAACTGGATTTCACACCTGCCCAACAATAGCTGGCTTCAATTCG 506
 QY 361 GCAGATGCTGGTTTTCAGCTGTGGATGGGGAACAGGAGGAAACGCTGTCTCGAAA 420
 Db 507 GCAGATGCTGGTTTTCAGCTGTGGATGGGGAACAGGAGGAAACGCTGTCTCGAAA 566
 QY 421 CACAGACACTCTCCATAGACCAAGATGATTTCTGGGCTTTTCAGTTATGATGAGATGGCT 480
 Db 567 CACAGACACTCTCCATAGACCAAGATGATTTCTGGGCTTTTCAGTTATGATGAGATGGCT 626
 QY 481 AGTTTGAACCTTCTCAGTGAATACTTTATTTTCAGAAAAACGGGCGAGGAAAGATC 540
 Db 627 AGTTTGAACCTTCTCAGTGAATACTTTATTTTCAGAAAAACGGGCGAGGAAAGATC 686
 QY 541 TATTATGTCGGCTATTTCAGAGGCAACCATGGGCTTTTATGCAATTTTCCACCATGCCA 600
 Db 687 TATTATGTCGGCTATTTCAGAGGCAACCATGGGCTTTTATGCAATTTTCCACCATGCCA 746
 QY 601 GAGCTGGCTCAGAAAAATCAAAATGTTATTTTGTAGCCCATAGCCACTGTTAAGCAT 660
 Db 747 GAGCTGGCTCAGAAAAATCAAAATGTTATTTTGTAGCCCATAGCCACTGTTAAGCAT 806
 QY 661 GCAAAAAGCCCGGACCAAAATTTTGTGTCGCCAGATATGATGATCAAGGGATTTGTT 720
 Db 807 GCAAAAAGCCCGGACCAAAATTTTGTGTCGCCAGATATGATGATCAAGGGATTTGTT 866

QY 721 GGCAGAGAGAGATTTCTGTATCAGACAGATTTCTCAGACAACTTGTATTTTACCTTCTG 780
 Db 867 GGCAGAGAGAGATTTCTGTATCAGACAGATTTCTCAGACAACTTGTATTTTACCTTCTG 926
 QY 781 GGCAGAGAGATTTCTGTATCAGATTTGTAGTAATATCATGTATTTCTTGGGTGGATTTCAAC 840
 Db 927 GGCAGAGATTTCTGTATCAGATTTGTAGTAATATCATGTATTTCTTGGGTGGATTTCAAC 986
 QY 841 ACCAACAATATGAACATGAGCCAGCAAGTGTATATGCTGCCACACTTCTTGGTGAACA 900
 Db 987 ACCAACAATATGAACATGAGCCAGCAAGTGTATATGCTGCCACACTTCTTGGTGAACA 1046
 QY 901 TCTGTGCAAAATATTTACACTGAGCCAGCAAGTGTATATGCTGCCACACTTCTTGGTGAACA 960
 Db 1047 TCTGTGCAAAATATTTACACTGAGCCAGCAAGTGTATATGCTGCCACACTTCTTGGTGAACA 1106
 QY 961 GACTGGGGAGTGAGACCAAAATCTTGGAAAAATGCAATCAGCCAACTCTCTGTAGGTAC 1020
 Db 1107 GACTGGGGAGTGAGACCAAAATCTTGGAAAAATGCAATCAGCCAACTCTCTGTAGGTAC 1166
 QY 1021 AGAGTCAGAGATATGAGCGTCCCTACAGCAATGTGGACAGAGGTGAGGCTTTTCA 1080
 Db 1167 AGAGTCAGAGATATGAGCGTCCCTACAGCAATGTGGACAGAGGTGAGGCTTTTCA 1226
 QY 1081 AATCCAGAGAGCGTGAATATGCTCTCTGTAGGTGACCAACCTCATCTACCAATAGAAAT 1140
 Db 1227 AATCCAGAGAGCGTGAATATGCTCTCTGTAGGTGACCAACCTCATCTACCAATAGAAAT 1286
 QY 1141 ATTCTGAATGGGCTACAGTGGATTTTCATCTGGGTGTTGGATGCTCTCAGCGTATGTAC 1200
 Db 1287 ATTCTGAATGGGCTACAGTGGATTTTCATCTGGGTGTTGGATGCTCTCAGCGTATGTAC 1346
 QY 1201 AATGAATATCATCTATGATGAGGAGGAGGAGCAACCTTTCCAGGAGCGTGTGAG 1260
 Db 1347 AATGAATATCATCTATGATGAGGAGGAGGAGCAACCTTTCCAGGAGCGTGTGAG 1406
 QY 1261 GCGGTATG 1269
 Db 1407 GCGGTATG 1415

RESULT 6

US-09-835-996A-9
 ; Sequence 9, Application US/09835996A
 ; Patent No. US20020142953A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ballinger, Dennis
 ; APPLICANT: Loeb, Debra
 ; APPLICANT: Montgomery, Julie
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Zhou, Ping
 ; APPLICANT: Goodrich, Ryle
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Zhao, Qing
 ; APPLICANT: Wehrman, Tom
 ; APPLICANT: Drmanac, Radoje
 ; APPLICANT: Ren, Feiyan
 ; APPLICANT: Qian, Xiahong
 ; APPLICANT: Wang, Dunrui
 ; TITLE OF INVENTION: MATERIALS AND METHODS RELATING TO LIPID METABOLISM
 ; FILE REFERENCE: 28110/35915A
 ; CURRENT APPLICATION NUMBER: US/09/835,996A
 ; PRIOR FILING DATE: 2001-04-15
 ; PRIOR APPLICATION NUMBER: US 60/197,137
 ; PRIOR FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: US 09/714,936
 ; PRIOR FILING DATE: 2000-11-17
 ; PRIOR APPLICATION NUMBER: US 09/667,298
 ; PRIOR FILING DATE: 2000-09-22
 ; PRIOR APPLICATION NUMBER: US 09/631,451
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: US 09/598,042

Query Match 97.4%; Score 1235.4; DB 10; Length 1384;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1236; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 ; PRIOR FILING DATE: 2000-06-20
 ; NUMBER OF SEQ ID NOS: 45
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 9
 ; LENGTH: 1384
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-835-996A-9
 QY 1 ATGTTGGAACCTTGTCAAGACAGTGGATTTCTCACAAGATGGAATGTGGCTTCTG 60
 Db 147 ATGTTGGAACCTTGTCAAGACAGTGGATTTCTCACAAGATGGAATGTGGCTTCTG 206
 QY 61 ATTCTGGTGGGTATATGTTCCAGAGAAATGTGAATTCAGTACATATGCCAATTAAGCT 120
 Db 207 ATTCTGGTGGGTATATGTTCCAGAGAAATGTGAATTCAGTACATATGCCAATTAAGCT 266
 QY 121 GTGGACCCAGAGCAATTCATGATATTAGTGAATCATCAACATCAAGCTATCCCTGT 180
 Db 267 GTGGACCCAGAGCAATTCATGATATTAGTGAATCATCAACATCAAGCTATCCCTGT 326
 QY 181 GAGGAATATGAAGTCGCAACTGAAAGATGGGTATATCCTTTCTGTAAACAGGATTCCTCGA 240
 Db 327 GAGGAATATGAAGTCGCAACTGAAAGATGGGTATATCCTTTCTGTAAACAGGATTCCTCGA 386
 QY 241 GGCCTAGTGCACCTAAGAGACAGAGTCCAGGCTGTGGTGTACTGACAGCATGGCCTA 300
 Db 387 GGCCTAGTGCACCTAAGAGACAGAGTCCAGGCTGTGGTGTACTGACAGCATGGCCTA 446
 QY 301 GTTGGAGTGTAGCAACTGGATTTCCAACTGCGCCCAACAAATAGCTTGGGCTTCATTCTG 360
 Db 447 GTTGGAGTGTAGCAACTGGATTTCCAACTGCGCCCAACAAATAGCTTGGGCTTCATTCTG 506
 QY 361 GCAGATGCTGTTTGTAGCTGTGGATGGGAAACAGCAGGGGAAACGGCTGGTCTGAAAA 420
 Db 507 GCAGATGCTGTTTGTAGCTGTGGATGGGAAACAGCAGGGGAAACGGCTGGTCTGAAAA 566
 QY 421 CACAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTCAGTTATGATGAGATGGCT 480
 Db 567 CACAGACACTCTCCATAGACCAAGATGAGTTCTGGGCTTTCAGTTATGATGAGATGGCT 626
 QY 481 AGGTTGACCTTCTGTGAGTGAATAAATTTATTTTTCAGAAAAACGGGCGCAGAAAAAGATC 540
 Db 627 AGGTTGACCTTCTGTGAGTGAATAAATTTATTTTTCAGAAAAACGGGCGCAGAAAAAGATC 686
 QY 541 TATTATGTCGGCTATTTCACAGGCAACCAACCATGGGCTTTTATGCAATTTCCACCATGCCA 600
 Db 687 TATTATGTCGGCTATTTCACAGGCAACCAACCATGGGCTTTTATGCAATTTCCACCATGCCA 746
 QY 601 GAGCTGGCTCAGAAAAATCAAAATGATTTTGTCTTAGCACCATAGCCATGTTTAAAGCAT 660
 Db 747 GAGCTGGCTCAGAAAAATCAAAATGATTTTGTCTTAGCACCATAGCCATGTTTAAAGCAT 806
 QY 661 GCAAAAACCCCGGAGCAAAATTTTGTCTGCAGATATGATGATCAAGGATTTCTTT 720
 Db 807 GCAAAAACCCCGGAGCAAAATTTTGTCTGCAGATATGATGATCAAGGATTTCTTT 866
 QY 721 GGCAAAAAGAAATTTCTGTATCAGACCAAGATTTCTCAGACAACTTGTATTACCTTGT 780
 Db 867 GGCAAAAAGAAATTTCTGTATCAGACCAAGATTTCTCAGACAACTTGTATTACCTTGT 926
 QY 781 GGCAGAGTGTATTTGTATCAGATTTGTAGTAATATCATGTATCTTGGGTGGATTTCAAC 840
 Db 927 GGCAGAGTGTATTTGTATCAGATTTGTAGTAATATCATGTATCTTGGGTGGATTTCAAC 986
 QY 841 ACCAACAATATGAACATGAGCCAGCAAGTGTATATGCTGCCACACTCTTGTCTGGAACA 900
 Db 987 ACCAACAATATGAACATGAGCCAGCAAGTGTATATGCTGCCACACTCTTGTCTGGAACA 1046

QY 901 TCTGTCAAAATATCTCACTGGAGCCAGGAGTGAATCTTGTGTGAATCTCGGGCAATTT 960
Db 1047 TCTGTCAAAATATCTCACTGGAGCCAGGAGTGAATCTTGTGTGAATCTCGGGCAATTT 1106
QY 961 GACTGGGGAGTGAGACCAAAATCTGGAATAATCAATCAGCCAACTCTCTGAAGTAC 1020
Db 1107 GACTGGGGAGTGAGACCAAAATCTGGAATAATCAATCAGCCAACTCTCTGAAGTAC 1166
QY 1021 AGAGTCAGAGATATGAGGTCCTTACAGCAATCTGACAGGAGGTCAGGACTGGCTTTCA 1080
Db 1167 AGAGTCAGAGATATGAGGTCCTTACAGCAATCTGACAGGAGGTCAGGACTGGCTTTCA 1226
QY 1081 AATCCAGAAGAGCTGAAAATCTGTCTCTGAGGTGACCAACCTCATCTACCATGAAGAT 1140
Db 1227 AATCCAGAAGAGCTGAAAATCTGTCTCTGAGGTGACCAACCTCATCTACCATGAAGAT 1286
QY 1141 ATTCTTGATGGGCTCAGTGAATTTTCATCTGGGGTTTGGATGCTCTCACCCTATGATAC 1200
Db 1287 ATTCTTGATGGGCTCAGTGAATTTTCATCTGGGGTTTGGATGCTCTCACCCTATGATAC 1346
QY 1201 AATGAATCATCATCTGATGACAGGAGGAGACCA 1237
Db 1347 AATGAATCATCATCTGATGACAGGAGGAGACCA 1383

RESULT 7
US-09-811-825-1
; Sequence 1, Application US/09811825
; Patent No. US2002014297A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES
; FILE REFERENCE: THEREOF
; FILE REFERENCE: CL001170
; CURRENT APPLICATION NUMBER: US/09/811,825
; CURRENT FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1206
; TYPE: DNA
; ORGANISM: Human
US-09-811-825-1

Query Match 88.4%; Score 1122; DB 10; Length 1206;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 148 AGTGAATCATCAACATCAAGGCTATCCCTGTGAGGATATCAAGTCCCACTGAAGAT 207
Db 76 AGTGAATCATCAACATCAAGGCTATCCCTGTGAGGATATCAAGTCCCACTGAAGAT 135
QY 208 GGGTATATCTTCTGTAAACAGGATCTCTCGAGGCTAGTGAACCAAGACAGGT 267
Db 136 GGGTATATCTTCTGTAAACAGGATCTCTCGAGGCTAGTGAACCAAGACAGGT 195
QY 268 TCCAGGCTGTGTGTATCTGACAGATGGCTAGTTGGAGGTCTAGCAACTGGATTTCC 327
Db 196 TCCAGGCTGTGTGTATCTGACAGATGGCTAGTTGGAGGTCTAGCAACTGGATTTCC 255
QY 328 AACCTGCCAACAAATAGCTGGCTTCATCTGGCAGATGCTGGTTTTCAGCTGGGATG 387
Db 256 AACCTGCCAACAAATAGCTGGCTTCATCTGGCAGATGCTGGTTTTCAGCTGGGATG 315
QY 388 GGGAAACAGCAGGGGAAACGCTGGTCTCGAAACCAAGACACTCTCCATAGACCAAGAT 447
Db 316 GGGAAACAGCAGGGGAAACGCTGGTCTCGAAACCAAGACACTCTCCATAGACCAAGAT 375
QY 448 GAGTCTTGGGCTTTTCAGTTATGATGATGGCTAGTGGTTTTCAGCTTCTCCTCAGTGAAC 507
Db 376 GAGTCTTGGGCTTTTCAGTTATGATGATGGCTAGTGGTTTTCAGCTTCTCCTCAGTGAAC 435

QY 508 TTTATTTTTCAGAAAACGGCCAGGAAAGATCTATTATGTGGCTATTTCACAGGGCACC 567
Db 436 TTTATTTTTCAGAAAACGGCCAGGAAAGATCTATTATGTGGCTATTTCACAGGGCACC 495
QY 568 ACCATGGGCTTTATTTGCATTTTCCACATGCTCCAGAGCTGGCTCAGAAAATCAAAATGTAT 627
Db 496 ACCATGGGCTTTATTTGCATTTTCCACATGCTCCAGAGCTGGCTCAGAAAATCAAAATGTAT 555
QY 628 TTTGCTTTTACACCCATAGCCTGTTTACCATGCAAAAAGCCCGGACCAAAATTTTG 687
Db 556 TTTGCTTTTACACCCATAGCCTGTTTACCATGCAAAAAGCCCGGACCAAAATTTTG 615
QY 688 TTGCTGCCAGATATGATGATCAAGGGATTTCTTTGGCAAAAAGAAATTTCTGTATCAGACC 747
Db 616 TTGCTGCCAGATATGATGATCAAGGGATTTCTTTGGCAAAAAGAAATTTCTGTATCAGACC 675
QY 748 AGATTCTTCAGACAACTGTTTATTTTACCTTTTGGCCAGGCTGATTCTGTATCAGATTGT 807
Db 676 AGATTCTTCAGACAACTGTTTATTTTACCTTTTGGCCAGGCTGATTCTGTATCAGATTGT 735
QY 808 AGTAATATCATGTTTACTTCTGGGTGATTCACACCAAAATATGAACATGAGCCGAGCA 867
Db 736 AGTAATATCATGTTTACTTCTGGGTGATTCACACCAAAATATGAACATGAGCCGAGCA 795
QY 868 AGTGTATATGCTGCCACACTCTTGTGGAAACATCTGTGCAAAATATTTTACACTGGAGC 927
Db 796 AGTGTATATGCTGCCACACTCTTGTGGAAACATCTGTGCAAAATATTTTACACTGGAGC 855
QY 928 CAGGCACTGATTTCTGGTGAATCTCCGGGCAATTTGACTGGGGAGTGAGACCAAAATCTG 987
Db 856 CAGGCACTGATTTCTGGTGAATCTCCGGGCAATTTGACTGGGGAGTGAGACCAAAATCTG 915
QY 988 GAAAAATGCAATCAGCAACTCTCTTAAGGTACAGAGTACAGAGATATGACGGTCCCTTACA 1047
Db 916 GAAAAATGCAATCAGCAACTCTCTTAAGGTACAGAGTACAGAGATATGACGGTCCCTTACA 975
QY 1048 GCAATGTGGACAGAGGTGAGGCTTCAATTCAGAGTACAGAGTACAGAGTATGAGCTGCTC 1107
Db 976 GCAATGTGGACAGAGGTGAGGCTTCAATTCAGAGTACAGAGTACAGAGTATGAGCTGCTC 1035
QY 1108 TCTGAGGTGACCAACTCTTACCAATAAGAAATATTTCTGAATGGGCTCACGTGGATTTT 1167
Db 1036 TCTGAGGTGACCAACTCTTACCAATAAGAAATATTTCTGAATGGGCTCACGTGGATTTT 1095
QY 1168 ATCTGGGGTTTGGATGCTCTCCTCAGGATGATGATCAAAATATCAATCTGTATGAGCAG 1227
Db 1096 ATCTGGGGTTTGGATGCTCTCCTCAGGATGATGATCAAAATATCAATCTGTATGAGCAG 1155
QY 1228 GAGGAGACCAACTTTTCCAGGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 1269
Db 1156 GAGGAGACCAACTTTTCCAGGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 1197

RESULT 8
US-09-880-107-3878
; Sequence 3878, Application US/09880107
; Patent No. US2002014298A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; CURRENT APPLICATION NUMBER: US/09/880,107
; FILE REFERENCE: 44921-5028-WO
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3878


```

; LENGTH: 2481
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 Z31690
US-09-880-107-3878

Query Match      41.6%; Score 527.8; DB 10; Length 2481;
Best Local Similarity 66.6%; Pred. No. 4.9e-162;
Matches 751; Conservative 0; Mismatches 376; Indels 0; Gaps 0;

Qy 115 AAAGCTGTGAGCCAGAGCAATTCATGAATATAGTGAATATCCTCAACATCAAGGCTAT 174
Db 112 ACAGCTGTGAGTCTCGAAGCAATTCATGAATATAGTGAATATCCTCAACATCAAGGCTAT 171

Qy 175 CCCTGTGAGGAATATGAGTCCGCACTGAAGATGGGTATATCCTTCTGTAAAGGAT 234
Db 172 CCTAGTGAAGGAATACCTAGTGTGAGCAGAGATGATATATCTGTGCCCTTAACCGAAT 231

Qy 235 CCTGAGGCTAGTGCAACCTTAAGAGACAGGTTCCAGGCTGTGGTGTACTGCGAGCAT 294
Db 232 CCTATGGAGGAGAACCACTTCGACAAAGTCCGAAACGAGTTGTTCTCTGCGAAT 291

Qy 295 GGCCTAGTGTGAGTGTAGCACTGGAATTTCCAACTGCGCCCAACAATAGCCTGGGCTTC 354
Db 292 GGCCTGTGCGAGATTTCTAGTAACTGGGTCAAAACCTTGCACACAGCAGCCTGGGCTTC 351

Qy 355 ATTCCTGGCAGATGCTGGTGTGAGTGTGAGTGGGCAACAGCAGGGAACGCTGGTCT 414
Db 352 ATTCCTGTGATGCTGGTGTGAGTGTGAGTGGGCAACAGCAGGGAATACCTGGTCT 411

Qy 415 CGAAACACAAGACACTCTCCATAGACCAAGATGAGTCTGGGCTTTTCAGTTATGATGAG 474
Db 412 CGAAACACAAGACACTCTCAGTTCTCAGGATGAATTTCTGGGCTTTTCAGTTATGATGAG 471

Qy 475 ATGGCTAGTGTGACCTTCTGCGAGTGAATTAATTTATTTGCGAGAAACGGGCGAGGAA 534
Db 472 ATGGCAAAATGACCTACCAAGCTTCCATTAATCTTCAATTAATTAATTAATTAATTAAT 531

Qy 535 AAGATCTATTATGTGGGTATTTACAGGGCAACCACTGGGCTTTATGCAATTTTCCACC 594
Db 532 CAAGTGATATTATGTGGGTATTTCTCAAGGCACCACTAGATTTTATGATATTTTACAG 591

Qy 595 ATGCCAGAGCTGGCTCAGAAATTCMAAATGATTTTGTCTTAGCAGCCATAGCCAGCTTT 654
Db 592 ATCCCTGAGCTGGCTTAAAGGATTAATATGTTTTTGGCCCTGGGCTCTGGGCTTCGCTC 651

Qy 655 AAGCATGCAAAAGCCCGGAGCAAAATTTTGTGCTGCCAGATATGATGATCAAGGGA 714
Db 652 GCCTTCTGTACTAGCCCTATGGCCAAATTAGGACGATTTACAGATCATCTCATTAAGAC 711

Qy 715 TTGTTTGGCAAAAGAAATTTCTGATCAGACCAAGATTTCTCAGACAACTTTGTTATTTAC 774
Db 712 TTATTTGGAGACAAAGAAATTTCTCCAGAGTGGTGTGTTTGAAGTGGCTGGGTACCCAC 771

Qy 775 CTTTGTGGCCAGGTGATTTCTGATCAGATTTGTAGTATATATCATGTTACTTCTGGTGG 834
Db 772 GTTTGACATCATGCTCATAGGAGCTCTGTGGAATCTCTGTTTCTTCTGTGTGGA 831

Qy 835 TTCAACACCAACATATGAAATGAGCGGAGCAAGTGTATATGTGTCGCCCACTCTTGCT 894
Db 832 TTTAATGAGAGAAATTTAAATATGTCTAGAGTGGATGTATATACACATTTCTCTGCT 891

Qy 895 GGAAACATCTGCAAAATTTCTCACTGAGGAGCAGGAGTGAATTTCTGTTGAATCTCGG 954
Db 892 GGAACCTCTGTGCAAAATTTCTCACTGAGGAGCAGGAGTGAATTTCTGTTGAATTTCAA 951

Qy 955 GCATTGTAGTGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCAACTCTGTTA 1014
Db 952 GCCTTTGACTGGGAGAGAGTCCCAAGAAATTTATTTTCAATACAGAGTATCTCTCC 1011

Qy 1015 AGGTACAGAGTCAGAGATATGACCGTCCCTACAGCAATGTGGACAGGAGGTGAGACTGG 1074

```

```

RESULT 9
US-10-198-846-13431
; Sequence 13431, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13431
; LENGTH: 2858
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 2, 2718, 2720, 2721, 2722, 2723, 2846, 2847, 2848,
; LOCATION: 2849, 2851, 2852, 2853, 2854, 2855, 2856, 2857, 2858
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-13431

```

```

Query Match      41.6%; Score 527.8; DB 15; Length 2858;
Best Local Similarity 66.6%; Pred. No. 5.4e-162;
Matches 751; Conservative 0; Mismatches 376; Indels 0; Gaps 0;

Qy 115 AAAGCTGTGAGCCAGAGCAATTCATGAATATAGTGAATATCCTCAACATCAAGGCTAT 174
Db 247 ACAGCTGTGAGTCTCGAAGCAATTCATGAATATAGTGAATATCCTCAACATCAAGGCTAT 306

Qy 175 CCCTGTGAGGAATATGAGTCCGCACTGAAGATGGGTATATCCTTCTGTAAAGGAT 234
Db 307 CCTAGTGAAGATACCTAGTTGAGACAGAGATGATATATCTGTGCTTTAACCGAAT 366

Qy 235 CCTGAGGCTAGTGCAACCTTAAGAGACAGGTTCCAGGCTGTGGTGTACTGAGCAT 294
Db 367 CCTCATGGGAGGAAGAACCAATTTCTGACAAAGGTCCCAACCAAGTTGTCTTCTGCAACAT 426

Qy 295 GGCCTAGTTGGAGGTGCTAGCAACTGGATTTTCCAACTGCCCAACAATAGCTGGGCTTC 354
Db 427 GGTCTGCTGGAGATTTCTAGTAACTGGGTCAAACTTGGCCACAGCAGGCTGGGCTTC 486

Qy 355 ATTCTGGCAGATGCTGGTGTGAGTGGGAAACAGCAGGGGAAACGCTGGTCT 414
Db 487 ATTCCTGTGATGCTGGTGTGAGTGGGAAACAGCAGGGGAAACATACCTGGTCT 546

Qy 415 CGAAACACAAGACACTCTCCATAGACCAAGATGAGTCTGGGCTTTTCAGTTATGATGAG 474
Db 547 CGAAACACAAGACACTCTCAGTTCTTCTCAGGATGAATTTCTGGGCTTTTCAGTTATGATGAG 606

```


QY 1075 CTTTCAATCCAGAGAGCTGAAATGCTCTCTCTGAGTGACCAACCTCTCTACCAT 1134
 Db |||||
 QY 5537 CTTGAGATGTACGAGCGCAATATCTTACTGATCAGATCCCACTTGGGTTCAT 5596
 Db |||||
 QY 1135 AAGAATATCTCTGAATGGCTCAGTGGAATTTCTCTGGGTTTGGATGCTCTCACCGT 1194
 Db |||||
 QY 5597 GAGAGCATTCGGAATGGGAGCATCTTGAATCTTCAATTTGGGCTGGATGCCCTTGGAG 5656
 Db |||||
 QY 1195 ATGTACATGAATCATCTCTGATGCAGCAGGAGGACCAACCT 1241
 Db |||||
 QY 5657 CTTTATAATAAATATTAATCTAATGAGGAATATCAGTGAAGCT 5703
 Db |||||

RESULT 11
 US-09-971-392-116
 ; Sequence 116, Application US/09971392
 ; Publication No. US20030134283A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Peterson, David P.
 ; APPLICANT: Pearson, Cecelia I.
 ; APPLICANT: Cocks, Benjamin G.
 ; TITLE OF INVENTION: GENES REGULATED IN DENDRITIC CELL DIFFERENTIATION
 ; FILE REFERENCE: PA-0029 US
 ; CURRENT APPLICATION NUMBER: US/09/971.392
 ; CURRENT FILING DATE: 2001-10-03
 ; PRIOR APPLICATION NUMBER: 60/237,652
 ; PRIOR FILING DATE: 2000-10-03
 ; NUMBER OF SEQ ID NOS: 260
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 116
 ; LENGTH: 2853
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Template ID: 410917.7
 ; NAME/KEY: unsure
 ; LOCATION: 207-395
 ; OTHER INFORMATION: a, t, c, g, or other
 US-09-971-392-116

Query Match 41.4%; Score 525.4; DB 13; Length 2853;
 Best Local Similarity 66.6%; Pred. No. 3.3e-161;
 Matches 751; Conservative 0; Mismatches 376; Indels 0; Gaps 0;
 QY 115 AAAGCTGTGGACCCAGAGCATTCATGAATATTAGTGAATCATCCAAATCAATCAAGGCTAT 174
 Db |||||
 QY 475 ACAGCTGTGGATCTCTGAACCAACCAATCATGATGTGAGTGAATATCTCTTACTGGGATTC 534
 Db |||||
 QY 175 CCCTGTGAGGATATGAGTGGCACTGAAGATGGGTATATCTCTTCTGTAAAGATTT 234
 Db |||||
 QY 535 CCTAGTGAGGAATACCTAGTTGAGACAGAGATGATATATCTGTGCCCTTAACCGAAT 594
 Db |||||
 QY 235 CCTGAGGCTTAGTGCAACCTTAAGAGACAGGTTCCAGGCTGTGGTGTCTTCTGACGAT 294
 Db |||||
 QY 595 CCTCATGGAGGAGAGACCATCTCTGACAAAGGTCCTCAACAGTTGTTCTCTGCAAT 654
 Db |||||
 QY 295 GGCTAGTTGAGGTGTAGCAATGATTTCCAACTGCTGCCCAACATAGACCTGGGCTTC 354
 Db |||||
 QY 655 GGCTTGTGGCAGATTTCTAGTAACTGGGTACAAACCTTCCCAACAGCAGCTGGGCTTC 714
 Db |||||
 QY 355 ATTCGTCAGATGCTGCTTTTGTAGCTGTGGATGGGAGCAGAGGGAACGCTGCTCT 414
 Db |||||
 QY 715 ATTTCTGTGATGCTGCTTTTGTAGCTGTGGATGGGAGCAGAGGGAATATCCTGGTCT 774
 Db |||||
 QY 415 CGAAAAACACAGACACTCTCAGTTCTCAGATGAATCTCTGGGCTTTTCAATTATGATGAG 474
 Db |||||
 QY 775 CGGAAACATAGACACTCTCAGTTCTCAGATGAATCTCTGGGCTTTTCAATTATGATGAG 834
 Db |||||
 QY 475 ATGCTAGGTTTGAACCTTCTCTGAGTGAATTAATTTTTCAGAAAAACGGGCGCAGGAA 534
 Db |||||
 QY 835 ATGGCAAAATATGACCTTACCAGCTTCCATTAACCTTCAATCTGAATAAACTGGCCCAAGAA 894
 Db |||||

QY 535 AAGATCTATTATGCTGGCTATTTCACAGGCGACCAACATGGCTTTATTGCACTTTCCACC 594
 Db |||||
 QY 895 CAAGTGTATTATGCTGGGTCTATTCTCAAGGCGACCACTATAGTTTTATTAGCATTTTCACAG 954
 Db |||||
 QY 595 ATGCCAGAGCTGGCTCAGAAAAATCAAAATGTAATTTTCTTTAGCACCCTAGCAGCTGTT 654
 Db |||||
 QY 955 ATCCCTGAGCTGGCTAAAAAGGATTAATTTTGGCTTGGCTCTGTTGGCTTCCGTC 1014
 Db |||||
 QY 655 AAGCATGCARAAAGCCCGGACCAATTTTGTGCTGCAGATATGATGATCAAGGGA 714
 Db |||||
 QY 1015 GCCTTCTGTACTAGCCCTATGCCCCAATTAGAGGATTAACAGATCATCTCAITTAAGGAC 1074
 Db |||||
 QY 715 TTGTTTGGCAAAAAGAAATTTCTGTATCAGACCAAGATTTCTCAGACCAACTTTTATTATAC 774
 Db |||||
 QY 1075 TTATTTGGAGACAAGAAATTTCTTCCCAGAGTGGCTTTTGAAGTGGCTGGGTACCCAC 1134
 Db |||||
 QY 775 CTTTGTGCCAGGTGATCTTGTATCAGATTTGTAGTAAATATCATGTTACTTCTGGGTGGA 834
 Db |||||
 QY 1135 GTTTGCACTCATGTACTGAAGGAGCTCTGTGGAATCTCTGTTTTCTTCTGTGTGGA 1194
 Db |||||
 QY 835 TTCAACCAACAATATGAACATGAGCGGAGCAAGTGTATATGCTGCCACACACTCTTGTCT 894
 Db |||||
 QY 1195 TTTAATGAGACAAATTTAATATGTTCTAGAGTGGATGATATATACACACATTTCTCTGCT 1254
 Db |||||
 QY 895 CGAACATCTGTCAAAAATATTCTACACTGGAGCGAGCAGTGAATTTCTGTGAACTCCGG 954
 Db |||||
 QY 1255 GGAACCTTCTGTGCAAAAACATGTTACACTGGAGCGAGGCTGTTAAATTTCCAAAAGTTCAA 1314
 Db |||||
 QY 955 GCATTTGACTGGGGGTGAGACCAAAATCTGAAAAATGCAATCAGCCAACTCCCTGTA 1014
 Db |||||
 QY 1315 GCCTTTGACTGGGGAAGCAGTGCCCAAGATTTATTTTCAACACAGAGTTATCTCTCCC 1374
 Db |||||
 QY 1015 AGGTACAGAGTCAGAGATATGACCGTCCCTACAGCAATTTGGGAGCAGGAGTCAAGACTGG 1074
 Db |||||
 QY 1375 ACATACATGTGAAGGACATCTTGTGCCGACTGCTGAGCGGGGTCACGACTGG 1434
 Db |||||
 QY 1075 CTTTCAATCCAGAGAGCAGTGAAAAATGCTGCTCTCTGAGGTGACCAACCTCACTACCAT 1134
 Db |||||
 QY 1435 CTTGAGATGCTACGACGTCATATCTTACTGACTCAGATCAGATCACCACCTTGGTGTCCAT 1494
 Db |||||
 QY 1135 AAGAAATTTCTGAATGGGCTCAGTGATTTCTATCTGGGTTTGGATGCTCTCTCACCGT 1194
 Db |||||
 QY 1495 GAGAGCATTCGGAATGGGAGCATCTTGACTTCAATTTGGGGCTGGATGCCCTTGGAGG 1554
 Db |||||
 QY 1195 ATGTACAAATGAATCATCTCATCTGATGAGGAGGAGACCAACCT 1241
 Db |||||
 QY 1555 CTTTATAATAAATTAATTAATCTAATGAGGAATATCAGTGAAGCT 1501
 Db |||||

RESULT 12
 US-10-108-260A-1262
 ; Sequence 1262, Application US/10108260A
 ; Publication No. US20040005560A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HELIX RESEARCH INSTITUTE
 ; TITLE OF INVENTION: No. US20040005560A1el full length cDNA
 ; FILE REFERENCE: H1-A0106
 ; CURRENT APPLICATION NUMBER: US/10/108,260A
 ; CURRENT FILING DATE: 2002-03-27
 ; NUMBER OF SEQ ID NOS: 5458
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 1262
 ; LENGTH: 2506
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-108-260A-1262

Query Match 36.4%; Score 462.2; DB 12; Length 2506;
 Best Local Similarity 67.0%; Pred. No. 1.9e-140;
 Matches 656; Conservative 0; Mismatches 323; Indels 0; Gaps 0;
 QY 263 CAGGTTCAGGCTGTGCTGTACTGACGATGCCCTAGTTGGAGGTGCTAGCAACTGGA 322
 Db |||||

Db 280 CTGGTCCCAACACAGTGTGCTCTGCTCAACATGGCTCGCTGCGCAGATCTAGTAACCTGGG 339
 Qy 323 TTTCGAACCTGCCCAACATAGCTGGGCTTCACTTCTGCGCAGATCTGCTGTTTGAAGTGT 382
 Db 340 TCACAAACCTTGGCAACACAGCCTGGGCTTCACTTCTGCTGATGCTGTTTGAAGTGT 399
 Qy 383 GGATGGGGAACAGCAGCGGGAACCGCTGCTCTCGAAGAACACAGACATCTCTCATAGACC 442
 Db 400 GGATGGGGAACAGCAGCGGGAACCTGCTCTCGAAGAACATAGACATCTCTAGTTTCTC 459
 Qy 443 AGATGAGTCTGGGCTTTCAGTATGATGAGATGCTAGGTTTGAACCTTCTCTGAGTGA 502
 Db 460 AGATGAGTCTGGGCTTTCAGTATGATGAGATGCTAGGTTTGAACCTTCTCTGAGTGA 519
 Qy 503 TAAACTTTATTTGTCAGAAACCGGCGCAGGAAAGATCTATTATGTCGGCTATTTCACAGG 562
 Db 520 TTAACCTTCTGATATAAACCTGGCCAGAACAGTGTATTATGTTGGTCACTTCTCAAG 579
 Qy 563 GCACCAACATGGGCTTTATGATTTTCCACCATGCCAGAGCTGGCTCAGAAAATCAAAA 622
 Db 580 GCACCAACATAGGTTTATAGCATTTTTCAGATCCCTGAGCTGGCTAAAGGATTAATA 639
 Qy 623 TGTATTTCTTTAGCACCATAGACCTGTTTAAAGCATGCAAAAAGCCCGGACCAAT 682
 Db 640 TGTATTTCTGGCTCTGCTGCTTCTGCTGCTTCTGCTGCTTCTGCTGCTTCTGCTGCT 699
 Qy 683 TTTTGTGCTGCGCATATGATGATCAAGGGATGTTTGGCAGAAAGATTTCTGTATC 742
 Db 700 TAGGACGATTACCATATCATCTCATTAAGGACTTATTGGACAGAAATTTCTTCCC 759
 Qy 743 AGACAGATTTCTCAGAACCTGTTTATTTACCTTTGGGCGAGTGAATCTTGTATCAGA 802
 Db 760 AGAGTGCCTTTTGAAGTGGCTGGGTACCCAGCTTTGCACTCATCTCATCTGAAGGAGC 819
 Qy 803 TTTGTAGTATATCATGTTTACTTCTGGGTGGATTTCAACACCAACATATGAACATGAGCC 862
 Db 820 TCTGTGGAATCTCTGTTTCTTCTGTTGGATTTAATGAGAGAAATTTAATATGTCTA 879
 Qy 863 GAGCAAGTGTATGCTGCGCACTCTTTGCTGGAACATCTTGTGCAAAATATTTACACT 922
 Db 880 GAGTGAATGTATATACACATCTCTGCTGGAACCTTCTGTCGCAAAACATGTTTACACT 939
 Qy 923 GGAGCCAGCAGTGAATCTGCTGACTCGGGCACTTTGACTGGGGGAGTGAGACCAAA 982
 Db 940 GGAGCCAGCAGTGTAAATTTCCAAAGTTTCAAGCTTTGACTGGGGAAGCAGTGCCAGA 999
 Qy 983 ATCTGGAATAATCAATCAGCCAACTCTCTGTAAGGTACAGAGTCAAGATATGACGGTCC 1042
 Db 1000 ATTATTTTCAATACACAGAGTATCTCTCCACGTAACATGTGAAGCAGATGCTTGTGC 1059
 Qy 1043 CTACAGCAATGTCAGAGAGTCTGAGTCTGCTTCAATCCAGAGACGTTGAATGTC 1102
 Db 1060 CCACTGCACTCTGGAGCGGGGTCAGACTGGCTTGCAGATGCTTACGACGTCATATCT 1119
 Qy 1103 TGCTCTCTGAGGTGACCAACCTCATCTACATAAGAAATTTCTGAAATGGGCTCAGTGG 1162
 Db 1120 TACTGACTCAGATCAACCACTTGGTGTTCATGAGAGCAATCCGGAATGGGAGCATCTTG 1179
 Qy 1163 ATTTCATCTGGGTTTGGATGCTCTTCAACCGTATGTAATGAATATTCATCTGATGC 1222
 Db 1180 ACTTCAATTTGGGCTTGGATGCTCTTCAACCGTATGTAATGAATATTCATCTGATGC 1239
 Qy 1223 AGCAGGAGGAGACCAACCT 1241
 Db 1240 GGAATATCAGTGAAGCT 1258

RESULT 13
 US-10-420-564-3
 ; Sequence 3, Application US/10420564
 ; Publication No. US20040001819A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bolen, Paul L.

; APPLICANT: Cihak, Paul, L.
 ; APPLICANT: Scharpf Jr., Lewis G.
 ; TITLE OF INVENTION: Recombinant Kid Pregastric Esterase and Methods for
 ; TITLE OF INVENTION: Its
 ; TITLE OF INVENTION: Production and Use
 ; FILE REFERENCE: IFP-0009
 ; CURRENT APPLICATION NUMBER: US/10/420,564
 ; CURRENT FILING DATE: 2003-04-22
 ; PRIOR APPLICATION NUMBER: US/10/043,665B
 ; PRIOR FILING DATE: 2002-09-24
 ; PRIOR APPLICATION NUMBER: US 09/186,489
 ; PRIOR FILING DATE: 1998-11-05
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 1411
 ; TYPE: DNA
 ; ORGANISM: Kid (Goat)
 ; US-10-420-564-3

Query Match 36.0%; Score 456.8; DB 12; Length 1411;
 Best Local Similarity 62.0%; Pred. No. 7.9e-139;
 Matches 741; Conservative 0; Mismatches 452; Indels 3; Gaps 1;
 Qy 42 AATGGAAATGCTGCTTCTGATTTCTGCTGGCTATATGTTCCAGAGAAATGTAATTCAGT 101
 Db 61 ATTGMAAATGCTGGCTACTTCTAAGCTGTGTTTCATCCACATCTCTGGAATGCATT 120
 Qy 102 ACATATGC---CAACTAAAGCTGTGACCCAGAGCAATTCATGAATATTAGTGAATCAT 158
 Db 121 TTGTTTCTTGGAAAAATTTGCTAAGAACCTTGAAGCCAGTATGAATGTGAGTCAGATGAT 180
 Qy 159 CCAACATCAAGGCTATCCCTGTGAGGAATATGAAGTCGCACTCGAAAGATGGGTATATCTCT 218
 Db 181 TTCTTCTGGGGCTACCCAGTGAGATGCATAAAGTTATACTGCAAGTGGCTATATCTCT 240
 Qy 219 TTCTGTTAAAGGATTCCTGAGGCTGTAGTGCACCTAAAGAGACAGGTTCCAGGCTGT 278
 Db 241 TCAGGTCTATCGGATTCCTCATGGAAGAAATGATGCTAATCATTTAGGTGAGAGACCTGT 300
 Qy 279 GGTGTTACTCAGCAGTGGCTAGTTCGAGGTGCTGCAACTGATTTCCAACTGCTCCAA 338
 Db 301 TGTGTTTCTGAGCATGCTTCTTCCCTCAGCTCAAACTGATTTCCAACTTCCAACTTCCAA 360
 Qy 339 CAATAGCTGGGCTTCAATTTCTGCGAGATGCTGTTTGTGAGCTGTGATGGGGAACAGCAG 398
 Db 361 CAACAGCTGGGCTTCTCTCTGCGAGATGCTGTTTATGAGCTGTGCTGGGGAACAGCAG 420
 Qy 399 GGGAAAGCCCTGCTCGAAGAACACAGACATCTCTCAGCAAGATGAGTCTGGGC 458
 Db 421 AGGAAACACTTGGGCGCAGGAACATTTATTAATTCACAGACTCTCCCTGAAATTTGGGC 480
 Qy 459 TTTCAAGTTATGATGAGTGGCTAGGTTTGAACCTTCTGCTGAGTGAATAAATTTATTTTGA 518
 Db 481 TTTCAAGTTTGAATGAATGGTGAATGACCTTCCATCTACAAATTTGATTTATCTTAA 540
 Qy 519 GAAAAAGGCGCCAGAAAGATCTATTATGTCGCTATTCAGAGGCAACCAATGGGCTT 578
 Db 541 GAGAACAGGACAGAGAGGCTACATATGTTGGCCATTCCTCCAGGCAACCAATGGGCTT 600
 Qy 579 TATTGCAATTTTCCACCATGCGAGCTGCTGCAAGAAATCAAAATGATTTTGTCTTTAGC 638
 Db 601 TGTGCGCTTTTCTACCAATCCCACTGGCTGAAAAATGGAAGTCTTCCATGCAATAGC 660
 Qy 639 ACCATAGCCACTGTTTAAAGCATGCAAAAAAGCCCGGACCAAAATTTTGTGTGTCAG 698
 Db 661 CCCAGTGGCCACAGTGAAGCAGACACCCAGAGCTGTTTAAACAACTTGCCTTATCTCA 720
 Qy 699 TATGATGATCAAGGATTTGTTGGCAAAAGAAATTTCTGTATCAGACAGATTTTCTAG 758
 Db 721 CTTCTCTTCAAGATTTATTTGGTAAACAAATGTTTCTACCCACACAAATTTTGTGAACA 780
 Qy 759 ACAACTGTTATTTTACCTTTGTCGCCCGGCTGATTTCTTGTATCAGATTTGTAGTAATCAT 818

Db 781 ATTTCTGGTGTGAAGTGTCTCTCTGTGAGACACTGATGTCTTTTGAAGTGCCTT 840
 Qy 819 GTTACTTCTGGTGGATTCAACACCAATAATGAAACATGAGCCGAGCAAGTGTATATGC 878
 Db 841 GTTTGCCAATCTGAGCTGACAAATAAAACTTCAACATGAGTGGCTTAGATGTATGT 900
 Qy 879 TGCCACACCTTCTGGAACATCTGCAAAATATTTCTACATCTGGAGCCAGCAGTGA 938
 Db 901 AGCACATAATCCAGCAGAGCTTCTGTTCAAAACATCTCTCACTGGAGACAGCTATTAA 960
 Qy 939 TTTCTGGTGAATCTCGGGCAATTTGACTGGGGAGTGAACCAAAATCTGGAAAAATGCCAA 998
 Db 961 GTCTGGGAAATTCGAAGCTTTGACTGGGAGCCTCAGTTGAGAACCTAATGATTATAA 1020
 Qy 999 TCAGCCAACTCTGTAGGTACAGAGTCAGAGATATGACGGTCCCTACAGCAATGGAC 1058
 Db 1021 TCAGCCCAACCTCCCACTCTACAATTTAAGAGCCATGAATGTCCCAATTCAGATGGAG 1080
 Qy 1059 AGGAGGTCAAGACTGGCTTTCAAAATCCAGAGACGTGAAAATGCTCTCTCAGGTGAC 1118
 Db 1081 TGCTGGCCAGACCTGTGGCTGACCTCAGGATGTTGACCTTTTCAAAACTCTC 1140
 Qy 1119 CAACCTCATCTACCAATGAATATTTCTGAAATGGGCTCAGTGAATTTCTCTGGGTTT 1178
 Db 1141 TAATCTCATTTCAACCAAGAAATTCCAAATTTACAATCATCTGGACTTTATCTGGGCAAT 1200
 Qy 1179 GGATGCTCTCAAGCTGTATCAATGAATCAATCATCTGATGAGCAGGAGGAGA 1234
 Db 1201 GGAATGACCTCAAGATTTCAATGAATTAATTTCTTTGATGCAAAAGACAAA 1256

RESULT 14

US-10-312-088-7

; Sequence 7, Application US/10312088

; Publication No. US20030219862A1

; GENERAL INFORMATION:

; APPLICANT: Agarwal, Pankaj

; APPLICANT: Cogswell, John P.

; APPLICANT: Kabnic, Karen S.

; APPLICANT: Lai, Ying-Ta

; APPLICANT: Martensen, Shelby A.

; APPLICANT: Murdoch, Paul R.

; APPLICANT: Smith, Randall F.

; APPLICANT: Strum, Jay C.

; APPLICANT: Xiang, Zhaoqing

; APPLICANT: Xie, Qing

; APPLICANT: Rizni, Safia K.

; TITLE OF INVENTION: NOVEL COMPOUNDS

; FILE REFERENCE: GPS0029

; CURRENT APPLICATION NUMBER: US/10/312,088

; CURRENT FILING DATE: 2002-12-20

; PRIOR APPLICATION NUMBER: PCT/US01/19929

; PRIOR FILING DATE: 2001-06-22

; PRIOR APPLICATION NUMBER: 60/213,161

; PRIOR FILING DATE: 2000-06-22

; PRIOR APPLICATION NUMBER: 60/213,156

; PRIOR FILING DATE: 2000-06-22

; NUMBER OF SEQ ID NOS: 44

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 7

; LENGTH: 1194

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-312-088-7

Query Match

Best Local Similarity 35.8%; Score 454.2; DB 13; Length 1194;

Matches 711; Conservative 0; Mismatches 428; Indels 0; Gaps 0;

Qy 86 GAAATGTGAATTCAGTACATATGCCAACTAAAGCTGTGGACCCAGAGCAATTCATGAATA 145

Db 44 GATCTATGTATGGTTATGACAGAAAGAAACAAATGCAAAACCTGAAGCTAATATGAATA 103

Qy 146 TTATGAAATCATCAACATCAAGGCTATCCCTGTGAGGAATATGAAGTGCACACTGAAG 205
 Db 104 TTAGCCAGATTATTTCTTACTGGGGTTATCCTTATGAAGAGTATGATGTTTACAAACAAAG 163
 Qy 206 ATGGTATATCTTTCTGTTTAAACAGAGATTCCTCAGGCGCTAGTGCACCTTAAGAAGACAG 265
 Db 164 ATGGTTATATCTTTGGAAATTTATAGGATTCACATGGAAGAGATGCCAGGGAGACAG 223
 Qy 266 GTTCAGGCCCTGTGGTGTACTGTGAGCATGGCCTAGTTGGAGGTGTGAGCACTGGATTTT 325
 Db 224 CTCCAAAGCCCTGCTGTGTATTTGAGCATGGCTTTAATTTGATCTGCCAGTAACCTGGATTT 283
 Qy 326 CCAACTCTCCCAACAAATAGACCTGGGCTTCATTCTGGCAGATGCTGTTTGTGACGTGTGA 385
 Db 284 GCACTCTCCCAACAAACAGTTTGGCTTTCTTCTGGCAGATAGTGGTTATGACGTGTGGT 343
 Qy 386 TGGGAAACAGCAGGGGAAACCGCTGTCTGAAAACAAAGACACTCTCCATAGACCAAG 445
 Db 344 TGGGAAACAGCGGAGGAAACACTTGGTCCAGAAACACCTTAAATTTGTACCGGAAATCAC 403
 Qy 446 ATGAGTTCTGGGCTTTCAGTTATGATGAGATGGCTGTGACCTTTCCTGCACTGATAA 505
 Db 404 CGGAATACTGGGCTTTCAGTTTGGATGAGATGGCTTAAATATGACCTTTCAGGCCAATCA 463
 Qy 506 ACTTTATTTTGCAGAAAACGGGCCAGGAAAGATCTATTATGTGCGCTATTTCACAGGGCA 565
 Db 464 ATTTTATCATAGAGAAACTGGACAGAGGCACTCTACTAGTGGGCACTTCACAGGCA 523
 Qy 566 CCACATAGGGCTTTATGCAATTTTCCACCTGCGAGCTGGCTCAGAAAATCAAATGT 625
 Db 524 CCACCATAGCTTTTATAGCATTCTTACAAACCCAGAACTGGCTTAAAGATTAAGATAT 583
 Qy 626 ATTTTGTCTTAGCACCCATAGCCACTGTTAAGCATGCAAAAAGCCCGGAGACCAATTTT 685
 Db 584 TTTTGTCACTGGCTCCAGTTGTCAAGTTAATACACCCAAAGTCTTATGAAAATCTAA 643
 Qy 686 TGTCTGCCAGATATGATGATCAAGGATTTGTTGGCAAAAAGAAATTTCTGTATCAGA 745
 Db 644 CAACCCCTTTCCAGGCGAGTAGTAAAGTGTGTTGGTGAACAAATGTTCCACCCTCAT 703
 Qy 746 CCAGATTTCTCAGACAACTGTTTATTTTACCTTTTGGCCAGGTGATTTCTTATCAGATTT 805
 Db 704 CATTTGTTGACCAATTTCAATGGCCCAAGTGTGCAATCGAAAGCTATTCGTGTATTT 763
 Qy 806 GTAGTAATATCATGTTACTTCTGGGTGGATTCAACACCAACAAATATGAACATGAGCCGAG 865
 Db 764 GCAGCAACTTCCCTATTTACTCTGAGTGGATTGATCCGCAAAACCTTAAATATGATCGCT 823
 Qy 866 CAAGTGTATATGCTGCCACACTCTTGTGGAACATCTGTGCAAAATATTTCTACACTGGA 925
 Db 824 TGGATGTTTATTTGTACACAACTCTCGGGGAAACATCTGTTCCAGAAATGCTGCACTGG 883
 Qy 926 GCCAGCAGTGAATTTCTGGTGAACCTCCGGGCAATTTGACTGGGGAGTGAACCAAAATC 985
 Db 884 CTCAGCTTTACCCTCTGATGATTTACAGACTTATGACTGGGGAAATGACCTGATAATA 943
 Qy 986 TGGAAAAATGCAATCAGCCAACTCTGTAGGTGACAGATGAGATGATGAGCGTCCCTA 1045
 Db 944 TGAACATTTACAATCAGAGTCTATCCCTCTATATATGACCTGACCTGCAATGAAGTGCCTA 1003
 Qy 1046 CAGCAATGAGCAGAGGTCAAGGCTGAGGCTTTTCAATCCAGAACGCTGAAAATGCTGC 1105
 Db 1004 CTGCTATTTGGGCTGGTGGACATGATGCTCTGTAAACACCCAGGATGTGGCCAGGATAC 1063
 Qy 1106 TCTCTGAGGTGACCAACCTCATCTACCAATAAGAAATTTCTGGAATGGGCTCAGCGTGAAT 1165
 Db 1064 TCCCTCAAATCAAGAGTCTTCAATTAAGCTTATGCGGAAATGGAACCACTTTGATTT 1123
 Qy 1166 TCACTGGGTTTGGATGCTCCTCCCTGATGCTACCAATGAATCATCCATCTGATGAG 1224
 Db 1124 TTGCTGGGCTCGATGCCCTCAACCGATGTACAGTGAATCATAGCTTTAATGAAG 1182

```

RESULT 15
US-10-420-564-4
; Sequence 4, Application US/10420564
; Publication No. US20040001819A1
; GENERAL INFORMATION:
; APPLICANT: Bolen, Paul L.
; APPLICANT: Cihak, Paul L.
; APPLICANT: Scharpf Jr., Lewis G.
; TITLE OF INVENTION: Recombinant Kid Pregastric Esterase and Methods for
; TITLE OF INVENTION: Its
; TITLE OF INVENTION: Production and Use
; FILE REFERENCE: IFF-0009
; CURRENT APPLICATION NUMBER: US/10/420,564
; CURRENT FILING DATE: 2003-04-22
; PRIOR APPLICATION NUMBER: US/10/043,665B
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 09/186,489
; PRIOR FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: bovine
US-10-420-564-4

Query Match          35.7%; Score 453.6; DB 12; Length 1134;
Best Local Similarity 62.7%; Pred. No. 7.6e-138;
Matches 705; Conservative 0; Mismatches 419; Indels 0; Gaps 0;

QY 111 AACTAAAGCTGTGGACCCAGAGCAATTCATGAATATTAGTGAATCATCCAAATCAAGG 170
DB 9 AAAAATTGCTAAGAACCCCTGAAGCCAGTAGTGAATGTTAGTCAGATGATTTCTACTGGGG 68

QY 171 CTATCCCTGTGGAGTATAGAGTCGCACTGAAGTGGTATATCTTTCTGTTAAAG 230
DB 69 CTACCCAGTGAAGTATGATTAAGTATTAAGTGGATGGTATATCTTCAAGGCTATCG 128

QY 231 GATTCCTCGAGGCTAGTGCACACCTTAAAGAACAGAGTTCACAGGCTGTGGTGTACTGCA 290
DB 129 GATTCCTCATGAAGAACAAATGCTAATCATTTAGGTAGAGACCTGTGTGTCTGCA 188

QY 291 GCATGCCCTAGTGTGGAGGTGTAGCACTGAGATTTCCAACTGCCCCAACATAGCCTGGG 350
DB 189 GCATGGTCTTCTTGGATCAGCCACAACTGGATTTCCAACTGCCCCAACAGAGCCTGGG 248

QY 351 CTTTCATCTGGCAGATGCTGTTTGGAGTGTGGATGGGACAGCAGGGAACGCTG 410
DB 249 CTTTCCTCTGGCAGATGCTGTTTGGAGTGTGGATGGGACAGCAGGGAACGCTG 308

QY 411 GTCTCGAAACACAGACACTCTCCATAGACCAAGATGAGTTCCTGGGCTTTTCAGTTATGA 470
DB 309 GGCCAGGAAACATTTATATCTATTACCACTCCCGGAAATCTGGGCTTTTCAGCTTTGA 368

QY 471 TGAGATGGCTAGTGTGACCTTCTCCAGTGTAACTTTATTTGCGAAGAACGGGCA 530
DB 369 TGAATGGCGGAATATGACCTTCCATCTACAAATGATTTTCACTTAAAGAGAACAGGCA 428

QY 531 GGAAAGATCTATTATGTGGGTATTACAGAGGACCACTGAGGCTTTATGCAATTTTC 590
DB 429 GAAGAAGCTACATATGTTGGCCATTTCCAGGACCACTGAGGCTTTATGCGCTTTTC 488

QY 591 CACCATGCGAGCTGGCTCAGAAATCAAAATGATTTGCTTTAGCAACCATAGCCAC 650
DB 489 TACCAGTCCCACTTGGCTGAAAAATCAAAAGTCTTCTATGCAATAGCCCCAGTTGCCAC 548

QY 651 TGTTAAGCATGCAAAAGCCCGGGACCAAAATTTTGTGCTGCGAGATATGATGATCAA 710
DB 549 AGTGAAGTACACCAAGAGCTGTTTAAACAACTTGCACTTATCTCTCTCTCTCAA 608

QY 711 GGGATTGTTGGCAAAAAAGAAATTTCTGTATCAGACCAAGATTTCTCAGAACCTTTAT 770

```

```

DB 609 GATTATATTTGGTCACAAAATGTTTACCCACACACTTTTTCGAAACAATTTCTTGGTGT 668
QY 771 TTACCTTTTGGCCAGGTGATTTCTGATCAGATTTGATGATATATCATGTTTACTTCTGGG 830
DB 669 TGAATGTGCTCCCGTGAGACACTGGATGTCCTTTGTGAAGATGCTTGTGTCCTATTAC 728
QY 831 TGGATTCAACACCAACAATATGAACATGAGCCGAGCAAGTGTATATGCTGCCCCACACTCT 890
DB 729 TGGATTGCAACATTAACAACTTCAACATGAGTCTGATGATGATATATAGCATATATCC 788
QY 891 TGTGGAACACTCTGTGCAAAATATTCTACACTGAGCCAGGAGTGAATTTCTGTTGAAC 950
DB 789 AGCAGGAACCTTCTGTTCAAAAACACCTCCACTGGAGACAGGCTGTTAAGTCTGGAAAT 848
QY 951 CCGGGCAATTTGACTGGGGAGTGAGACCAAAATCTGGAATAATGCAATCAGCCACTCC 1010
DB 849 CCAAGCTTTTGGACTGGGGAGCCCCATATCAAGAACCTAATGCAATATCATCAGCCAC 908
QY 1011 TGTAAAGGTACAGAGTACAGAGATATGAGGTCCTTACAGCAATGTGGACAGGAGTCA 1070
DB 909 TCCCATCTACAATTTAAACAGCCCATGATGTCCTCAATTTGAGTGTGAGTGTGCAATGA 968
QY 1071 CTGGCTTTCAATCCAGAGACAGTGAATGCTGCTCTCTGAGGTGACCAACCTCATCTA 1130
DB 969 CTTGTTGGCTGACCCCTCAGGATGTTGACTTTCTGCTTTCAAAAACCTCTCTAATCT 1028
QY 1131 CCATAAGATATTTCTGAAATGGGCTCAGCTGGATTTTCTCTGGGTTTGGATGCTCTCA 1190
DB 1029 CCACAGGAATTTCCAAATTTACAACTCACTTGGACTTTTCTGGGCAATGGATGCACT 1088
QY 1191 CCGTATGTACAATGAATCAATCCATCTGATGACAGAGGAGGAGA 1234
DB 1089 AGAAGTTTACAATGAATTTGTTCTTTGATGGCCGAAGACAAAA 1132

```

Search completed: February 19, 2004, 14:12:32
Job time : 463.878 secs